

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 6016.11 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataacatatattcatc.....cgttgtgcgtgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2032	100.0	2032	6	AR257469	AR257469 Sequence
2	2032	100.0	2032	6	AX148742	AX148742 Sequence
3	2032	100.0	2032	8	ARQARAH	L34402 Arachis hyp
4	2032	100.0	2041	6	AR257470	AR257470 Sequence
5	2019.2	99.4	2032	6	AX155331	AX155331 Sequence
6	1930	95.0	1930	6	BD172109	BD172109 Peanut al
7	1930	95.0	1930	6	AR257463	AR257463 Sequence
8	1930	95.0	1930	6	BD107898	BD107898 Methods a
9	1805.2	88.8	1978	6	AX155332	AX155332 Sequence
10	1752.4	86.2	1949	6	AR257467	AR257467 Sequence
11	1752.4	86.2	1949	8	ARQARAH	L38853 Arachis hyp
12	1608	79.1	2674	8	AF432231	AF432231 Arachis h
13	1395.4	68.7	2332	8	AY581852	AY581852 Arachis h
14	1375.4	67.7	1418	8	AY581851	AY581851 Arachis h
15	954.4	47.0	1035	8	AY581850	AY581850 Arachis h
16	891.2	43.9	1047	8	AY581849	AY581849 Arachis h
17	750	36.9	750	6	AR257466	AR257466 Sequence
18	468.6	23.1	1257	8	LCU51424	AJ551424 Lens culi
19	465.8	22.9	1433	8	PSVIC	Y00722 Pisum sativ

20	461	22.7	1561	8	VFVIC	Y00462 Vicia faba
21	455.4	22.4	1248	8	LCU551425	AJ551425 Lens culi
22	449	22.1	1248	8	AJ626897	AJ626897 Pisum sat
23	449	22.1	1596	8	PSVICK	X67429 Pisum sativ
24	446	21.9	1791	8	AY500372	AY500372 Lupinus a
25	444.2	21.9	1248	8	AJ626898	AJ626898 Pisum sat
26	435.4	21.4	1548	8	VNVICLN	Z71987 V.narbonens
27	408.4	20.1	1251	8	AB008679	AB008679 Glycine m
28	406.8	20.0	1320	6	BD175903	BD175903 Regulatio
29	406.8	20.0	1320	6	AR202568	AR202568 Sequence
30	406.8	20.0	1320	6	AR486455	AR486455 Sequence
31	404	19.9	1632	8	AB008678	AB008678 Glycine m
32	404	19.9	1818	6	BD175901	BD175901 Regulatio
33	404	19.9	1818	6	AR202566	AR202566 Sequence
34	404	19.9	1818	6	AR486453	AR486453 Sequence
35	403.6	19.9	1955	8	GMASBC	X17698 Soybean mRN
36	403.6	19.9	1477	8	AB030840	AB030840 Glycine m
37	403.6	19.9	1477	8	AB030841	AB030841 Glycine m
38	402	19.8	1680	8	AB008680	AB008680 Glycine m
39	402	19.8	1884	8	AB113351	AB113351 Glycine m
40	400.8	19.7	2070	8	AB030839	AB030839 Glycine m
41	398.8	19.6	2042	8	AB030838	AB030838 Glycine m
42	384	18.9	1336	8	AY221105	AY221105 Glycine m
43	378.8	18.6	2069	8	PSA276875	AJ276875 Pisum sat
44	347.8	17.1	1858	8	VNCONVN	Z71986 V.narbonens
45	339.4	16.7	1494	8	GMU14006	U14006 Glycine max

ALIGNMENTS

RESULT 1
LOCUS AR257469 2032 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 21 from patent US 6486311.
ACCESSION AR257469
VERSION AR257469.1 GI:27307482
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2032)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannon,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 21 26-NOV-2002;
FEATURES Location/Qualifiers
source 1..2032
/organism="unknown"
/mol_type="genomic DNA"

Query Match	100.0%	Score 2032;	DB 6;	Length 2032;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGAG	60	
DB	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGAG	60	
QY	61	GGTTTCTCCAGTCGCTGTGCTAGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120	
DB	61	GGTTTCTCCAGTCGCTGTGCTAGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120	
QY	121	TGCCAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCCGCAGAGGTGCCTCCA	180	
DB	121	TGCCAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCCGCAGAGGTGCCTCCA	180	
QY	181	GAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTTCGCTGCAACCA	240	
DB	181	GAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTTCGCTGCAACCA	240	
QY	241	GCTCGNGATGATGCTCTCGTGTGTCTATGATCCTCGAGGACACACTGGGCCACCAACCA	300	

Db 241 |||||GCTCGAGTATGATCCTCGTGTGCTCTATGATCCTCGAGGACACACTGGCACCAACCA 300
Qy 301 |||||ACGTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCGGAGACTACGATGATGACCG 360
Db 301 |||||ACGTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCGGAGACTACGATGATGACCG 360
Qy 361 |||||CGCTCAACCCCGAAGAGAGAGAGGAGCGCGATGGGACACAGCTGGACCGAGGAGCGTGA 420
Db 361 |||||CGCTCAACCCCGAAGAGAGAGAGGAGCGCGATGGGACACAGCTGGACCGAGGAGCGTGA 420
Qy 421 |||||AAGAGAAGAGACTGGAGACAAACCAAGAGAAGATTGGAGCGACCAAGTCAACAGAGCC 480
Db 421 |||||AAGAGAAGAGACTGGAGACAAACCAAGAGAAGATTGGAGCGACCAAGTCAACAGAGCC 480
Qy 481 |||||ACGGAAATAAGGCCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 |||||ACGGAAATAAGGCCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 |||||TGTGAGGAGAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Db 541 |||||TGTGAGGAGAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Qy 601 |||||CGCTACGGGAACCAAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAG 660
Db 601 |||||CGCTACGGGAACCAAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAG 660
Qy 661 |||||GCAGTTTCAGATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACTCT 720
Db 661 |||||GCAGTTTCAGATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACTCT 720
Qy 721 |||||TGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCCAGCAAGGCGACAGCCAC 780
Db 721 |||||TGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCCAGCAAGGCGACAGCCAC 780
Qy 781 |||||CGTGACCGTAGCAAAATGGAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840
Db 781 |||||CGTGACCGTAGCAAAATGGAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840
Qy 841 |||||CAGATCCCATCCGGTTTCAATCTTCTACATCTTGACCCCGATGACCAACAGAACTCTAG 900
Db 841 |||||CAGATCCCATCCGGTTTCAATCTTCTACATCTTGACCCCGATGACCAACAGAACTCTAG 900
Qy 901 |||||AGTAGTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC 960
Db 901 |||||AGTAGTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC 960
Qy 961 |||||GAGCAGCCGAGACCAATCATCTTAATTGAGGCTTTCAGCAGGAATACGTTGGAGGCGGC 1020
Db 961 |||||GAGCAGCCGAGACCAATCATCTTAATTGAGGCTTTCAGCAGGAATACGTTGGAGGCGGC 1020
Qy 1021 |||||CTTCAATGCGGAATTCATATGAGATCGGAGGCTGCTGTAGAGAGAAATGCGAGGAGTGA 1080
Db 1021 |||||CTTCAATGCGGAATTCATATGAGATCGGAGGCTGCTGTAGAGAGAAATGCGAGGAGTGA 1080
Qy 1081 |||||GCAAGAGGAGAGAGGCGCAGAGCGATGAGTACTCGAGTACTGAGAACATGAGAGGAGT 1140
Db 1081 |||||GCAAGAGGAGAGAGGCGCAGAGCGATGAGTACTCGAGTACTGAGAACATGAGAGGAGT 1140
Qy 1141 |||||GATAGTCAAAGTGTCAAAGAGACGTTGAAGAACTTACTAAGCAGCGTTAAATCCGTTCTC 1200
Db 1141 |||||GATAGTCAAAGTGTCAAAGAGACGTTGAAGAACTTACTAAGCAGCGTTAAATCCGTTCTC 1200
Qy 1201 |||||AAAGAAAGGCTCCGAGAGAGGAGATATCACCACCAATCACTTGAAGAGGCGGA 1260
Db 1201 |||||AAAGAAAGGCTCCGAGAGAGGAGATATCACCACCAATCACTTGAAGAGAGGCGGA 1260
Qy 1261 |||||GCCGATCTTTCTAACTTTGGGAAGTTATTGAGGTGAAGCCAGACAAAGAAACCC 1320
Db 1261 |||||GCCGATCTTTCTAACTTTGGGAAGTTATTGAGGTGAAGCCAGACAAAGAAACCC 1320
Qy 1321 |||||CGAGTTTCAGGACCTGGAATGATGCTACCTGTGTAGAGATCAAGAAAGGAGCTTTGAT 1380

Db 1321 |||||CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGAT 1380
Qy 1381 |||||GCTCCCACTTCAACTCAAGGCCATGGTTATCTGTCGTCAACAAAGGAACCTGGAAA 1440
Db 1381 |||||GCTCCCACTTCAACTCAAGGCCATGGTTATCTGTCGTCAACAAAGGAACCTGGAAA 1440
Qy 1441 |||||CCTTGAATCTGCTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACCGCGGGAAGAGAGA 1500
Db 1441 |||||CCTTGAATCTGCTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACCGCGGGAAGAGAGA 1500
Qy 1501 |||||GGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCAGGTT 1560
Db 1501 |||||GGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCAGGTT 1560
Qy 1561 |||||GAAGGAAGCGATGTTTCATCAGCCAGAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Db 1561 |||||GAAGGAAGCGATGTTTCATCAGCCAGAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Qy 1621 |||||CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGAACTCTTCTTGC 1680
Db 1621 |||||CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGAACTCTTCTTGC 1680
Qy 1681 |||||AGGTGATAGGACAATGTGTAGACACAGATAGAGAAAGCAAGCAAGGATTTAGCAATCCC 1740
Db 1681 |||||AGGTGATAGGACAATGTGTAGACACAGATAGAGAAAGCAAGCAAGGATTTAGCAATCCC 1740
Qy 1741 |||||TGGGTCCGGTGACCAAGTTGAGAGCTCATCAAAACCAAGAGGATCTCACTTTGTGAG 1800
Db 1741 |||||TGGGTCCGGTGACCAAGTTGAGAGCTCATCAAAACCAAGAGGATCTCACTTTGTGAG 1800
Qy 1801 |||||TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCCTTGAGAA 1860
Db 1801 |||||TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCCTTGAGAA 1860
Qy 1861 |||||AGAGATCAAGAGGAGGAAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAGGC 1920
Db 1861 |||||AGAGATCAAGAGGAGGAAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAGGC 1920
Qy 1921 |||||TTTTAACTGAGAACTGAGGCACTTGTATGATGATTAAGATCAAGCTTTTGTACT 1980
Db 1921 |||||TTTTAACTGAGAACTGAGGCACTTGTATGATGATTAAGATCAAGCTTTTGTACT 1980
Qy 1981 |||||CTACTATCAAAAACCTTATCAATAAATAAAGCTTTGCGTTGTTCTCC 2032
Db 1981 |||||CTACTATCAAAAACCTTATCAATAAATAAAGCTTTGCGTTGTTCTCC 2032

RESULT 2
AXI48742
LOCUS AXI48742 2032 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 5 from Patent WO0136621.
ACCESSION AXI48742
VERSION AXI48742.1 GI:14347296
KEYWORDS
SOURCE
ORGANISM
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1
REFERENCE
AUTHORS Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
TITLE Down-regulation and silencing of allergen genes in transgenic
peanut seeds
JOURNAL Patent: WO 0136621-A 5 25-MAY-2001;
Alabama A & M University (US)
FEATURES
source location/Qualifiers
1. 2032
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"

ORIGIN

Query Match				Score 2032; DB 6; Length 2032;			
Best Local Similarity				100.0%; Pred. No. 0;			
Matches 2032; Conservative				0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60				
Db	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60				
Qy	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTCTGCTTCAGTTTCTGGAACGCA	120				
Db	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTCTGCTTCAGTTTCTGCAACGCA	120				
Qy	121	TGCCAAGTCATCACTTTACAGAGAAACACAGAGAACCCCTGCGCCACAGAGGTGCCTCCA	180				
Db	121	TGCCAAGTCATCACTTTACAGAGAAACACAGAGAACCCCTGCGCCACAGAGGTGCCTCCA	180				
Qy	181	GAGTTGTCAAACAGAACCCGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCCACCA	240				
Db	181	GAGTTGTCAAACAGAACCCGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCCACCA	240				
Qy	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCTGAGGACACACTGGGACACCAACCA	300				
Db	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCTGAGGACACACTGGGACACCAACCA	300				
Qy	301	ACGTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTAAGCATGATGACCG	360				
Db	301	ACGTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTAAGCATGATGACCG	360				
Qy	361	CCGTCAAACCCGAGAGGAGGAGCCGATGGGACACAGCTGGACCCGAGGGAGCGTGA	420				
Db	361	CCGTCAAACCCGAGAGGAGGAGCCGATGGGACACAGCTGGACCCGAGGGAGCGTGA	420				
Qy	421	AAGAGAGAGACTGGAGACAAACAGAGAAAGATTGGAGCGGACCAAGTCAATCAGAGCC	480				
Db	421	AAGAGAGAGACTGGAGACAAACAGAGAAAGATTGGAGCGGACCAAGTCAATCAGAGCC	480				
Qy	481	ACGGAATAAAGCCCGAG	540				
Db	481	ACGGAATAAAGCCCGAG	540				
Qy	541	TGTGAGGAGAAACATCTCGGAACCAACCCCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC	600				
Db	541	TGTGAGGAGAAACATCTCGGAACCAACCCCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC	600				
Qy	601	CCGCTAGCGGAACCAAAACGGTAGGATCCGGGTCCTGACAGAGTTTGAACAAAGGTCAG	660				
Db	601	CCGCTAGCGGAACCAAAACGGTAGGATCCGGGTCCTGACAGAGTTTGAACAAAGGTCAG	660				
Qy	661	GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT	720				
Db	661	GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT	720				
Qy	721	TGTTCTTCCAGCAGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCCAC	780				
Db	721	TGTTCTTCCAGCAGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCCAC	780				
Qy	781	CGTGACGCTAGCAATAGGCAATACAGAAAGAGCTTTAATCTTGAAGGGCCATGCACT	840				
Db	781	CGTGACGCTAGCAATAGGCAATACAGAAAGAGCTTTAATCTTGAAGGGCCATGCACT	840				
Qy	841	CAGAAATCCCATCCGTTTCAATTTCTTACATCTTGAACCGCCATGACAAACAGAACTCAG	900				
Db	841	CAGAAATCCCATCCGTTTCAATTTCTTACATCTTGAACCGCCATGACAAACAGAACTCAG	900				
Qy	901	AGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGGCGAGTTTGAAGATTTCTTCCCGGC	960				
Db	901	AGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGGCGAGTTTGAAGATTTCTTCCCGGC	960				
Qy	961	GAGCAGCCGAGACCAATCATCTTCTAGAGGCTTCAGCAGGAATACGTTGGAGGCCGC	1020				
Db	961	GAGCAGCCGAGACCAATCATCTTCTAGAGGCTTCAGCAGGAATACGTTGGAGGCCGC	1020				
Qy	1021	CTTCAATGCGGAATTCATAGAGATACGAGGGTGTCTGTTAGAGAGAGATGACGAGGTCGA	1080				

RESULT 3
ARQARAH
LOCUS
DEFINITION
ACCESSION

ARQARAH
Arachis hypogaea (clone P41b) Ara h I mRNA, complete cds.
2032 bp mRNA linear PLN 24-MAY-1996
L34402

1021 CTTCAATGCGGAATTCATAGAGATACGAGGGTGTCTGTTAGAGAGAGATGACGAGGTGA 1080
1081 GCAAGAGGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT 1140
1081 GCAAGAGGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT 1140
1141 GATAGTCAAGAGTCAAGAGGACGTTGAAGAACTTACTTAAGCACGCTAAATCCGCTCTC 1200
1141 GATAGTCAAGAGTCAAGAGGACGTTGAAGAACTTACTTAAGCACGCTAAATCCGCTCTC 1200
1201 AAAGAAGGCTCCGAAGAAGAGAGGAGATATCACAAACCAATCACTTGAAGAGAGCGGA 1260
1201 AAAGAAGGCTCCGAAGAAGAGAGGAGATATCACAAACCAATCACTTGAAGAGAGCGGA 1260
1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAGAACCC 1320
1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAGAACCC 1320
1321 CCAGCTTCAGGACCTGGACATGATGCTCACTCTGCTAGAGATCAAGAGAGGAGCTTTGAT 1380
1321 CCAGCTTCAGGACCTGGACATGATGCTCACTCTGCTAGAGATCAAGAGAGGAGCTTTGAT 1380
1381 GCTCCCACTTTCAACTCAAAAGGCCATGTTATCGTCTGCTCAACAAAGGAACTGGA 1440
1381 GCTCCCACTTTCAACTCAAAAGGCCATGTTATCGTCTGCTCAACAAAGGAACTGGA 1440
1441 CTTTGAATCTGCTGCTGTTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAGAGAGGA 1500
1441 CTTTGAATCTGCTGCTGTTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAGAGAGGA 1500
1501 GGACGAGAGCAAGTT 1560
1501 GGACGAGAGCAAGTT 1560
1561 GAAGGAGCGGATGCTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
1561 GAAGGAGCGGATGCTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
1621 CGAATCTCAATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAGAGATCTTCTTGC 1680
1621 CGAATCTCAATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAGAGATCTTCTTGC 1680
1681 AGGTGATAGGACAAATGTGATAGACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGTTTCC 1740
1681 AGGTGATAGGACAAATGTGATAGACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGTTTCC 1740
1741 TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGAGAGAGAGAGAGAGAG 1800
1741 TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGAGAGAGAGAGAGAGAG 1800
1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCTGAGAA 1860
1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCTGAGAA 1860
1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1921 TTTTAACTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
1921 TTTTAACTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
1981 CTACTATCAAAAACCTTATCAATAAATAAAGGTTTGTGCGTGTGTTCTTCTCC 2032
1981 CTACTATCAAAAACCTTATCAATAAATAAAGGTTTGTGCGTGTGTTCTTCTCC 2032

VERSION	L34402.1	GI:602435	
KEYWORDS	allergen.		
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
REFERENCE	1 (bases 1 to 2032)		
AUTHORS	Burks,A.W., Cockrell,G., Stanley,J.S., Helm,R.M. and Bannon,G.A.		
TITLE	Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity		
JOURNAL	J. Clin. Invest. 96 (4), 1715-1721 (1995)		
MEDLINE	96013631		
PUBMED	7560062		
COMMENT	On Dec 16, 1994 this sequence version replaced gi:508640. Original source text: Arachis hypogaea (strain Florunner).		
FEATURES	Location/Qualifiers		
source	1..2032		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/strain="Florunner"		
	/db_xref="taxon:3818"		
	/clone="F41b"		
	/dev_stage="seed"		
Gene	1..2032		
	/gene="Ara h I"		
5'UTR	<1..49		
	/gene="Ara h I"		
CDS	50..1930		
	/gene="Ara h I"		
	/codon_start=1		
	/protein_id="AAB00861.1"		
	/db_xref="GI:602436"		
	/translation="MRGRVSLMLLILVLASVATHAKSSPYOKKTENPCARCLQ SCQPPDDLKQKACSRCTKLEYDFRCVTDPRGHTGTTNQRPSPGERTGRQFYDD DRQPRREGGWSGAPREREREDWQREDWRPSPHQPKRIPREGEQEWGT PSQHVRETSRNPFYFSPRRFSTRYGNQGRIRVLQRFSQFQNLQNRHIVQIE AKPNTLVLPHKADADNIIYIOQQTATVTVANGNRKSNFLDRGHALRIPSGFISVLN RDHONLRVAKLSMPVNTQGFQDPFPASSRQSSVLOGFSRNTLEAAFNAEFNIRR VLLENAGGEQERQORRWSSTRSENNEGVIVKSKHVEBELTKHAKSVSKGSEEG DITPNILREGPSLUSNFGKLFVKPQKNFQQLQDMLTCVSEIKGALMLPHFNS KAMWIVVNNKGJLEJELVAVRKEQQORREEDDEBEEGNSREVRRYRLKKEGD VFIMPAHFVAIINASSELHLGFGINAENNHRIFLAGDKNDVIDIEKQAKDLAPPGS GSQVEKLIKNQESHFVSARPOSQSPPSPKESPEKEDQBEENQGGKPLLSILKA FN"		
3'UTR	1931..2032		
	/gene="Ara h I"		
polyA_signal	2005..2010		
	/gene="Ara h I"		
polyA_site	2032		
	/gene="Ara h I"		
ORIGIN			
Query Match	100.0%;	Score 2032;	DB 8; Length 2032;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATCAGAGGGAG	60
Db	1	AATAATCATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATCAGAGGGAG	60
QY	61	GGTTTCTCCACTGATGCTGTGCTAGGGATCCCTGTGCTGGCTTCAGTTTCTGCAACGCA	120
Db	61	GGTTTCTCCACTGATGCTGTGCTAGGGATCCCTGTGCTGGCTTCAGTTTCTGCAACGCA	120
QY	121	TGCCAAGTCATCACCTTACAGAGAAACACAGAAACCCCTCGCCACAGAGTGCTTCCA	180
Db	121	TGCCAAGTCATCACCTTACAGAGAAACACAGAAACCCCTCGCCACAGAGTGCTTCCA	180
QY	181	GAGTTGTCAACAGGACCGGATGACTTTGAAGCAAAAGGCGATCGAGTCTCGCTGCACCAA	240
Db	181	GAGTTGTCAACAGGACCGGATGACTTTGAAGCAAAAGGCGATCGAGTCTCGCTGCACCAA	240

QY	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCAACCA	300
Db	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCAACCA	300
QY	301	ACGTTCCCTCCTCAGGGGAGCGGACACGTTGGCCGCGCAACCCGGAGACTACGATGATGAC	360
Db	301	ACGTTCCCTCCTCAGGGGAGCGGACACGTTGGCCGCGCAACCCGGAGACTACGATGATGAC	360
QY	361	CGGTCAACCCCAAGAGAGGAGGCGGATGGGGACCGCTGACCGAGGGAGGCGTCA	420
Db	361	CGGTCAACCCCAAGAGAGGAGGCGGATGGGGACCGCTGACCGAGGGAGGCGTCA	420
QY	421	AAGAGAAGAAGACTCGAGACCAACCAAGAGAAATTGGAGGCGCAACCAAGTCAATCAGCAG	480
Db	421	AAGAGAAGAAGACTCGAGACCAACCAAGAGAAATTGGAGGCGCAACCAAGTCAATCAGCAG	480
QY	481	ACGGAATAAAGGCCGCGAAGAGAGAAAGAGGAGAAACAGAGTGGGGAAACACCAAGT	540
Db	481	ACGGAATAAAGGCCGCGAAGAGAGAAAGAGGAGAAACAGAGTGGGGAAACACCAAGT	540
QY	541	TGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGCGGTTTAGCAC	600
Db	541	TGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGCGGTTTAGCAC	600
QY	601	CCGCTACGGGAACCAAAAACGGTAGGATCCGGGTCCTGCGAGAGTTTGAACCAAGTCAAG	660
Db	601	CCGCTACGGGAACCAAAAACGGTAGGATCCGGGTCCTGCGAGAGTTTGAACCAAGTCAAG	660
QY	661	GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACTCT	720
Db	661	GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACTCT	720
QY	721	TGTTCTTCCCAAGCACGCTGATGATAAATCTGTTATCCAGAGGCGGCAAGCCAC	780
Db	721	TGTTCTTCCCAAGCACGCTGATGATAAATCTGTTATCCAGAGGCGGCAAGCCAC	780
QY	781	CCTGACCGCTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGAAGGGGCGCATGCACT	840
Db	781	CCTGACCGCTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGAAGGGGCGCATGCACT	840
QY	841	CAGAAATCCCATCCGGTTTCATTTCCCTAATCTTGAACCGCCATGACAACCAAGAACTCAG	900
Db	841	CAGAAATCCCATCCGGTTTCATTTCCCTAATCTTGAACCGCCATGACAACCAAGAACTCAG	900
QY	901	AGTAGCTAAATCTCCATGCGCGTTTAAACACCCCGCCAGTTTGAAGGATTTCTTCCCGCG	960
Db	901	AGTAGCTAAATCTCCATGCGCGTTTAAACACCCCGCCAGTTTGAAGGATTTCTTCCCGCG	960
QY	961	GAGCAGCGGAGACCAATCATCTTTCAGGGGCTTCAGCAGGAAATAGCTTGGAGGCGCG	1020
Db	961	GAGCAGCGGAGACCAATCATCTTTCAGGGGCTTCAGCAGGAAATAGCTTGGAGGCGCG	1020
QY	1021	CTTCAATCGGAAATTCATAGATACGAGGGTGTCTGTAGAGAAATGACGAGGAGTGA	1080
Db	1021	CTTCAATCGGAAATTCATAGATACGAGGGTGTCTGTAGAGAAATGACGAGGAGTGA	1080
QY	1081	GCAAGAGGAGAGGCGCAGAGCGATGAGTACTCGGAGTAGTAGAGAAATGAAGGAGT	1140
Db	1081	GCAAGAGGAGAGGCGCAGAGCGATGAGTACTCGGAGTAGTAGAGAAATGAAGGAGT	1140
QY	1141	GATAGTCAAGTGTCAAAGGAGCAGTTTGAAGAACTTACTAAGCACGCTTAAATCCGCTCTC	1200
Db	1141	GATAGTCAAGTGTCAAAGGAGCAGTTTGAAGAACTTACTAAGCACGCTTAAATCCGCTCTC	1200
QY	1201	AAAGAAAGCGCTCCGAAAGAGGAGGAGATATCACCAACCAATCACTTTCAGAGAGGGA	1260
Db	1201	AAAGAAAGCGCTCCGAAAGAGGAGGAGATATCACCAACCAATCACTTTCAGAGAGGGA	1260
QY	1261	GCCCGATCTTTCTAAACAACTTTGGGAAGTATTTTGGGTGAAGCCAGACAAGAAACCC	1320
Db	1261	GCCCGATCTTTCTAAACAACTTTGGGAAGTATTTTGGGTGAAGCCAGACAAGAAACCC	1320

QY	1321	CCAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAAAGGAGCTTTGAT	1380
Db	1321		
QY	1381	GCTCCACACATTTCAACTCAAAAGGCCATGGTTATCGTCGTCTCAACAAAGGAACTCGAAA	1440
Db	1381		
QY	1441	CCTTGAACTCGTGCTGTAAGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGGA	1500
Db	1441		
QY	1501	GGACGAGACGAAGAGAGAGGAGTAACAGAGAGGTCGTAGGTACACAGCGAGGTT	1560
Db	1501		
QY	1561	GAAGGAGGCGATGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACCGCTTCCTC	1620
Db	1561		
QY	1621	CGAATCCATCTCGTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCCTTGC	1680
Db	1621		
QY	1681	AGTGATAGGACAAATGTGATAGACAGATAGAGAGCAAGCAAGATTTAGCATCC	1740
Db	1681		
QY	1741	TGGGTGGGTGAAACAAGTTGAGAGGCTCATCAAAAACCAAGGAATCTCACCTTGTGAG	1800
Db	1741		
QY	1801	TGCTGCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCCTGAGAA	1860
Db	1801		
QY	1861	AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCTCTTCAATTTTGAAGGC	1920
Db	1861		
QY	1921	TTTTAACTGAGATGGAGGCAACTTGTATGATCGATTAAGATCAAGCTTTTGTACT	1980
Db	1921		
QY	1981	CTACTATCCAAAACCTTATCAATAAATAAAGCTTTGTGCGTTGTTCTCC	2032
Db	1981		
RESULT 4			
AR257470			
LOCUS	2041 bp	DNA	linear
DEFINITION	Sequence 23 from patent US 6486311.		PAT 20-DEC-2002
ACCESSION	AR257470		
VERSION	AR257470.1	GI:27307483	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2041)		
AUTHORS	Burke,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,		
	Helm,R.M. and Bannon,G.A.		
TITLE	Peanut allergens and methods		
JOURNAL	Patent: US 6486311-A 23 26-NOV-2002;		
FEATURES	Location/Qualifiers		
source	1..2041		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	100.0%;	Score 2032;	DB 6; Length 2041;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

QY	1	AATAATCATATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGGAG	60
Db	1		
QY	61	GGTTTCTCCACTGATGCTGCTAGGGAATCCTTGTCTCTGGCTTCAAGTTTCTGCAACGCA	120
Db	61		
QY	121	TGCCAAGTCATCACCTTTACAGAGAAACACAGAGAAACCCCTGCGCCACAGAGGTGCTCCA	180
Db	121		
QY	181	GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGATCTCGCTGCACCAA	240
Db	181		
QY	241	GCTCGAGTATGATCCTCGTGTGCTCTATGATCCTCGAGGACACACTGGGCAACCAACCA	300
Db	241		
QY	301	ACGTTCCCTCCAGGGGAGCGGACACGTCGGCGGCCAAACCCGGAGACTACGATGATGACCG	360
Db	301		
QY	361	CGGTCAACCCGAAAGAGAGGAGCCGATGGGGACCCAGCTGGGACCCAGTGGACCCGAGGAGCGTGA	420
Db	361		
QY	421	AGAGGAAGAGACTCGAGACACCAAGAGAAGATTGGAGGCGACCAAGTCAATCAGCAGCC	480
Db	421		
QY	481	ACGGAATAAGCGCCCGAAGAGAGAGAGAAACAAAGAGTGGGAAACACCAAGGTAGCCA	540
Db	481		
QY	541	TGTGAGGAGAGAAACATCTCGGAAACAAACCTTTCTACTTCCCGTCAAGCGGTTTAGCAC	600
Db	541		
QY	601	CGCTACGGGAACCAAAACCGGTAGGATCCGGGTCTGCGAGAGGTTTGACCAAGGTCAAG	660
Db	601		
QY	661	GCAGTTTCAGAAATCTCCAGAAATCAAGTATTCGAGATCGAGGCGCAACCTTAACACTCT	720
Db	661		
QY	721	TGTTCTTCCCAAGCAGCTGATGATTAACATCTTGTGTTATCCAGCAAGGGCAAGCCAC	780
Db	721		
QY	781	CGTGACCGTAGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT	840
Db	781		
QY	841	CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACAAACAGAACTCAG	900
Db	841		
QY	901	AGTAGCTAAAATCTCCATGCGCTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC	960
Db	901		
QY	961	GAGCAGCGGAGACCAATCATCTTCTGAGGCTTCAGCAGGAATACCTTGAGGCGCC	1020
Db	961		
QY	1021	CTTCAATCGGAATTCATGAGATACGAGGCGTGTGTTAGAGAAATGACAGGAGTGA	1080
Db	1021		

```
QY 1081 GCAAGAGGAGAGAGGCGAGAGCGATGGAGTACTCGGAGTAGTGAGAAACAATCAAGGAGT 1140
Db 1081 GCAAGAGGAGAGAGGCGAGAGCGATGGAGTACTCGGAGTAGTGAGAAACAATCAAGGAGT 1140
QY 1141 GATAGTCAAGTGTCAAGAGGACAGTTGAAGAACTTACTAAGCAGCGTAAATCGTCTC 1200
Db 1141 GATAGTCAAGTGTCAAGAGGACAGTTGAAGAACTTACTAAGCAGCGTAAATCGTCTC 1200
QY 1201 AAAGAAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTCTGAGAGAAGGCGA 1260
Db 1201 AAAGAAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTCTGAGAGAAGGCGA 1260
QY 1261 GCCCGATCTTTTCAACAACTTTGGGAAGTTATTTGAGGTGAAGCCACAGCAAGAAGAACCC 1320
Db 1261 GCCCGATCTTTTCAACAACTTTGGGAAGTTATTTGAGGTGAAGCCACAGCAAGAAGAACCC 1320
QY 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGAT 1380
QY 1381 GGTCCACACACTTCAACTCAAAAGGCCATGGTTATCGTCTGTCGTCACAAAGGAACTGGAAA 1440
Db 1381 GGTCCACACACTTCAACTCAAAAGGCCATGGTTATCGTCTGTCGTCACAAAGGAACTGGAAA 1440
QY 1441 CTTGAACTCGTGGCTGTGAAGAAAGAGCAACAACAGAGGGGACCGCGGAAGAGAGGA 1500
Db 1441 CTTGAACTCGTGGCTGTGAAGAAAGAGCAACAACAGAGGGGACCGCGGAAGAGAGGA 1500
QY 1501 GGNCGAAGCAGAGAGAGGAGGAGTAAACAGAGAGGTTCGTAGGTACACAGCGAGGTT 1560
Db 1501 GGNCGAAGCAGAGAGAGGAGGAGTAAACAGAGAGGTTCGTAGGTACACAGCGAGGTT 1560
QY 1561 GAAAGGAGCGGATGTGTTCAATCATGCGCAGCGCTCATCCAGTAGGCATCAACGCTTCCTC 1620
Db 1561 GAAAGGAGCGGATGTGTTCAATCATGCGCAGCGCTCATCCAGTAGGCATCAACGCTTCCTC 1620
QY 1621 CGAATCTCAATCTGCTTGGCTTCGGTATCAACCGTGA AAAACAACAGAGATCTTCTTGC 1680
Db 1621 CGAATCTCAATCTGCTTGGCTTCGGTATCAACCGTGA AAAACAACAGAGATCTTCTTGC 1680
QY 1681 AGGTGATAGGACNATGTATAGACAGATAGAGAGCAAGCGAAGGATTTAGCATTC 1740
Db 1681 AGGTGATAGGACNATGTATAGACAGATAGAGAGCAAGCGAAGGATTTAGCATTC 1740
QY 1741 TGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAACACAGAAAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAACACAGAAAGGAATCTCACTTTGTGAG 1800
QY 1801 TGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTGAGAA 1860
QY 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCACCTCCCTTCAATTTGAGGCG 1920
Db 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCACCTCCCTTCAATTTGAGGCG 1920
QY 1921 TTTTAACTCAGAAATGAGGCAACTTGTATGTATCGATAAAGATCAAGCTTTTGTACT 1980
Db 1921 TTTTAACTCAGAAATGAGGCAACTTGTATGTATCGATAAAGATCAAGCTTTTGTACT 1980
QY 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032
Db 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032
```

```
RESULT 5
AX155331
LOCUS AX155331
DEFINITION Sequence 4 from Patent WO0140264.
ACCESSION AX155331
VERSION AX155331.1 GI:14536766
KEYWORDS Arachis hypogaea (peanut)
```

```
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1
REFERENCE 1
AUTHORS Bannon,G.A., Burks,W.A., Caplan,M.J., Sampson,H. and Sosin,H.
TITLE Peptide antigens
JOURNAL Patent: WO 0140264-A 4 07-JUN-2001;
The University of Arkansas (US)
; MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK
(US)
FEATURES
source
1..2032
/molecule="Arachis hypogaea"
/db_xref="taxon:3818"
ORIGIN
Query Match 99.4%; Score 2019.2; DB 6; Length 2032;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2024; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 AATAATCATATATATATTCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
Db 1 AATAATCATATATATATTCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
QY 61 GGTTCCTCCATCGATGCTTGTCTAGGGATCCTTGTCTTGGCTTCAGTTCTTGCACCGA 120
Db 61 GGTTCCTCCATCGATGCTTGTCTAGGGATCCTTGTCTTGGCTTCAGTTCTTGCACCGA 120
QY 121 TGCCAAAGTCATCACCTTACCAAGAGAAAACAGAGAAACCCCTCGCCGAGAGTGCTCCA 180
Db 121 GGCCAAGTCATCACCTTACCAAGAGAAAACAGAGAAACCCCTCGCCGAGAGTGCTCCA 180
QY 181 GAGTTGTCAAACAGGAACCGGATGACTTGAAGCAAAGGATGCGAGTCTCGCTGCACCAA 240
Db 181 GAGTTGTCAAACAGGAACCGGATGACTTGAAGCAAAGGATGCGAGTCTCGCTGCACCAA 240
QY 241 GCTCGAGTATGATCTCGTGTCTATGATCTCGAGGACACACTGGCGCCACCAACCA 300
Db 241 GCTCGAGTATGATCTCGTGTCTATGATCTCGAGGACACACTGGCGCCACCAACCA 300
QY 301 ACGTTCCCTCTCAGGGAGCGGACACAGTGGCGCCCAACCCGAGAGACTACGATGACCG 360
Db 301 ACGTTCCCTCTCAGGGAGCGGACACAGTGGCGCCCAACCCGAGAGACTACGATGACCG 360
QY 361 CGGTCAACCCGAGAGAGGAAGGCGCGATGGGGACAGCTGGACCGAGGAGCGTGA 420
Db 361 CGGTCAACCCGAGAGAGGAAGGCGCGATGGGGACAGCTGGACCGAGGAGCGTGA 420
QY 421 AAGAGAAGAAGTGGAGACAACCAAGAGAAGATTGGAGGCGACCAAGTCTACAGCGCC 480
Db 421 AAGAGAAGAAGTGGAGACAACCAAGAGAAGATTGGAGGCGACCAAGTCTACAGCGCC 480
QY 481 ACGGAAAAAAGGCGCGGAGGAAGAGAGAAAGAGTGGGGAACACCAAGGTAGCCA 540
Db 481 ACGGAAAAAAGGCGCGGAGGAAGAGAGAAAGAGTGGGGAACACCAAGGTAGCCA 540
QY 541 TGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCGTCGAAGCGGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCGTCGAAGCGGTTTAGCAC 600
QY 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
Db 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
QY 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAACACCTTAACTCT 720
Db 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAACACCTTAACTCT 720
QY 721 TGTTCCTCCGAGCAGCTGATGCTGATAACATCTCTTGTATTCAGGAGGCAAGCCAC 780
Db 721 TGTTCCTCCGAGCAGCTGATGCTGATAACATCTCTTGTATTCAGGAGGCAAGCCAC 780
```

Db 721 TGTTCCTCCACGACGCTGATGCTGATAACATCCTTGTATTATCCAGCAGGGCGAAGCCAC 780
Qy 781 CGTGACCGTAGAAATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAGGGCCATGCACT 840
Db 781 CGTGACCGTAGAAATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAGGGCCATGCACT 840
Qy 841 CAGAAATCCCAATCCGGTTTCAATTTCTTACATCTTTGAACCGCCATGACAAACAGAACCTCAG 900
Db 841 CAGAAATCCCAATCCGGTTTCAATTTCTTACATCTTTGAACCGCCATGACAAACAGAACCTCAG 900
Qy 901 AGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGGCAGTTTGGAGATTTCTTCCCGGC 960
Db 901 AGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGGCAGTTTGGAGATTTCTTCCCGGC 960
Qy 961 GAGCAGCCGAGACAATCATCTACTTTCAGAGGCTTCAGCAGGAATACGTTGGAGGCCG 1020
Db 961 GAGCAGCCGAGACAATCATCTACTTTCAGAGGATTCAGCAGGAATACGTTGGAGGCCG 1020
Qy 1021 CTTCAATCGCGAATTCATAGATACGAGGGGTCTGTTTGAAGAGAAATGACAGAGGTGA 1080
Db 1021 CTTCAATCGCGAATTCATAGATACGAGGGGTCTGTTTGAAGAGAAATGACAGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGGCAGAGGGATGGAGTACTCGAGGTAGTGAGAAACAATGAAGGAGT 1140
Db 1081 GCAAGAGGAGAGGGCAGAGGGATGGAGTACTCGAGGTAGTGAGAAACAATGAAGGAGT 1140
Qy 1141 GATAGTCAAGAGTCAAGAGGACGTTTGAAGAACTTCTAAGCAGCGCTTAAATCCGTCTC 1200
Db 1141 GATAGTCAAGAGTCAAGAGGACGTTTGAAGAACTTCTAAGCAGCGCTTAAATCCGTCTC 1200
Qy 1201 AAAGAAAGGCTCCGAGAGAGGGAGATATCACCAACCCCAATCACTTGAGAGAGGGCA 1260
Db 1201 AAAGAAAGGCTCCGAGAGAGGGAGATATCACCAACCCCAATCACTTGAGAGAGGGCA 1260
Qy 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAAAGAACCC 1320
Db 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAAAGAACCC 1320
Qy 1321 CCAGCTTCAGGACCTGAGACATGATGCTCACCTGTAGAGATCAAAAGAGGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGAGACATGATGCTCACCTGTAGAGATCAAAAGAGGAGCTTTGAT 1380
Qy 1381 GCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGCTCAACAAAGAGACTCGAA 1440
Db 1381 GCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGCTCAACAAAGAGAACTCGAA 1440
Qy 1441 CCTTGAACCTCGTGGCTGTAGAAAAGAGCAACACAGAGGGGACGGCGGAGAGAGGA 1500
Db 1441 CCTTGAACCTCGTAGCTGTAGAAAAGAGCAACACAGAGGGGACGGCGGAGAGAGGA 1500
Qy 1501 GGAAGAGAGAGAGAGAGAGAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
Db 1501 GGAAGAGAGAGAGAGAGAGAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
Qy 1561 GAAGGAGGGGATGTGTTCAATCATGCCAGCAGCTCCAGTAGCCATCAACGCTTCCTC 1620
Db 1561 GAAGGAGGGGATGTGTTCAATCATGCCAGCAGCTCCAGTAGCCATCAACGCTTCCTC 1620
Qy 1621 CGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTGC 1680
Db 1621 CGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTGC 1680
Qy 1681 AGGTGATAGGACAAATGTGATAGACAGATAGAGAAAGCAAGCGAATTTAGCATTCCTC 1740
Db 1681 AGGTGATAGGACAAATGTGATAGACAGATAGAGAAAGCAAGCGAATTTAGCATTCCTC 1740
Qy 1741 TGGGTGGGTGAACAATTTGAGAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTGGGTGAACAATTTGAGAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGTCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT

Qy 1861 AGAGGATCAAGAGGAGGAAAAACAAGGAGGGAAGGTCACCTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAAAAACAAGGAGGGAAGGTCACCTCTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTCAGAAATGGAGCAACTTGTATGATCGATAATAAGATCAAGCTTTTGTACT 1980
Db 1921 TTTTAACTCAGAAATGGAGCAACTTGTATGATCGATAATAAGATCAAGCTTTTGTACT 1980
Qy 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032
Db 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032

RESULT 6

BD172109
LOCUS Peanut allergen and method. 1930 bp DNA linear PAT 18-FEB-2003
DEFINITION
ACCESSION BD172109
VERSION BD172109.1 GI:28413407
KEYWORDS JP 2002223783-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1930)
AUTHORS Jr,W.A.B., Helm,R.M., Cockrell,G., Steven, Stanley,J. and Bannon,G.A.

TITLE

Peanut allergen and method
Patent: JP 200223783-A 2 13-AUG-2002;
UNIVERSITY OF ARKANSAS

JOURNAL

OS Unidentified
PN JP 200223783-A/2

COMMENT

PF 21-NOV-2001 JP 2001356754
PR 29-DEC-1995 US 60/009.455,04-MAR-1996 US 08/610.424 PI
WESLEY A BURKS JR,RICKI M HELM,GAEI COCKRELL,STEVEN PI J
STANLEY,GARY A BANNON
PC C12N15/09,A61K38/00,A61P37/04,C07K14/415,C12N15/00,A61K37/02
CC Strandedness: Double;
CC Topology: Linear;
CC Identified as Ara h I cDNA

PH KEY

Location/Qualifiers
FT source 1..1930
FT /organism='Unidentified'.

FEATURES

source
1..1930
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 95.0%; Score 1930; DB 6; Length 1930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATAATCATATATATTCATCATCATCTATATAGTAGTAGCAGGAGCAATGAGGGAG 60
Db 1 AATAATCATATATATTCATCATCATCTATATAGTAGTAGCAGGAGCAATGAGGGAG 60
Qy 61 GGTTCCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA 120
Db 61 GGTTCCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA 120
Qy 121 TGCCAAGTTCATCCTTACCAAGAAAAACAGAGAAACCCCTCGGCCAGAGGTGCTTCCA 180
Db 121 TGCCAAGTTCATCCTTACCAAGAAAAACAGAGAAACCCCTCGGCCAGAGGTGCTTCCA 180
Qy 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGATCTCGCTGCAACCA 240
Db 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGATCTCGCTGCAACCA 240
Qy 241 GCTCGAGTATGATCTCGTTGTGTCTATGATCTCTGAGGACACACTGGGCACCAACCA 300

241	Db	GCTCGAGTATGATCCTCGTTGTGTTATGATCTCTCGAGGACACACTGGCACCAACCAACCA	300
301	Qy	ACGTTTCCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCCGGAGACTACGATGATGACCG	360
301	Db		
301	Db	ACGTTTCCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCCGGAGACTACGATGATGACCG	360
361	Qy	CCGTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACCCAGCTGGACCGAGGGAGCGGTGA	420
361	Db	CCGTCAACCCCGAAGAGAGGAGGAGGCGCGATGGGACCCAGCTGGACCGAGGGAGCGGTGA	420
421	Qy	AAGAGAGAGAGACTGGGAGACCAACCAAGAGAAGATTGGAGGGACCAAGTCAATCAGAGACC	480
421	Db		
421	Db	AAGAGAGAGAGACTGGGAGACCAACCAAGAGAAGATTGGAGGGACCAAGTCAATCAGAGACC	480
481	Qy	ACGGAATAATAGGCCGAGGAGAGAGAGAGAAAGAGTGGGGAAACACCAAGGTAGGCCA	540
481	Db		
481	Db	ACGGAATAATAGGCCGAGGAGAGAGAGAGAAAGAGTGGGGAAACACCAAGGTAGGCCA	540
541	Qy	TGTGAGGGAAGAAACATCTCGGAAACAACCCCTTCTACTTCCCGTCAAGGCGTTTAGCAC	600
541	Db	TGTGAGGGAAGAAACATCTCGGAAACAACCCCTTCTACTTCCCGTCAAGGCGTTTAGCAC	600
601	Qy	CCGCTAGCGGAAACAAACCGGTAGGATCCGGGTCCTGAGAGGTTTGACCAAGGTCAG	660
601	Db	CCGCTAGCGGAAACAAACCGGTAGGATCCGGGTCCTGAGAGGTTTGACCAAGGTCAG	660
661	Qy	GCAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT	720
661	Db		
661	Db	GCAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT	720
721	Qy	TGTTCTTTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTCCAGCAAGGCGCAAGCCAC	780
721	Db		
721	Db	TGTTCTTTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTCCAGCAAGGCGCAAGCCAC	780
781	Qy	CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAACTCTTGACAGAGGCGCAATGCACT	840
781	Db	CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAACTCTTGACAGAGGCGCAATGCACT	840
841	Qy	CAGAACTCCCAATCCGGTTTCAATTCCTACATCTTGAAACCGGCATGACAAACAGAACCTCAG	900
841	Db		
841	Db	CAGAACTCCCAATCCGGTTTCAATTCCTACATCTTGAAACCGGCATGACAAACAGAACCTCAG	900
901	Qy	AGTAGCTAAATCTCCATGCGCGTTAAACACACCCGCGCAGTTTGAGGATTTCTTCCCGCG	960
901	Db		
901	Db	AGTAGCTAAATCTCCATGCGCGTTAAACACACCCGCGCAGTTTGAGGATTTCTTCCCGCG	960
961	Qy	GAGCAGCGGAGACCAATCACTACTTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCGCG	1020
961	Db	GAGCAGCGGAGACCAATCACTACTTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCGCG	1020
1021	Qy	CTTCAATGCGGAATTTCAATGAGATACCGAGGGTGCTGTTAGACAGAAATGCGAGGAGTGA	1080
1021	Db		
1021	Db	CTTCAATGCGGAATTTCAATGAGATACCGAGGGTGCTGTTAGACAGAAATGCGAGGAGTGA	1080
1081	Qy	GCAAGAGAGAGAGGCGCAGAGGCGATCGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT	1140
1081	Db		
1081	Db	GCAAGAGAGAGAGGCGCAGAGGCGATCGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT	1140
1141	Qy	GATAGTCAAAAGTGTCAAAGGAGCAGCTGTGAAGAACTTTACTAAGCACCGCTAAATCCGTC	1200
1141	Db	GATAGTCAAAAGTGTCAAAGGAGCAGCTGTGAAGAACTTTACTAAGCACCGCTAAATCCGTC	1200
1201	Qy	AAAGAAAGGCTCCGAAGAGAGGAGGATATCACCAACCCCAATCAACTTTGAGAGNAGCGGA	1260
1201	Db	AAAGAAAGGCTCCGAAGAGAGGAGGATATCACCAACCCCAATCAACTTTGAGAGNAGCGGA	1260
1261	Qy	GCCCGATCTTCTAAACAACTTTTGGGAAGTTATTGAGGTGAAGCCGACCAAGAGAACC	1320
1261	Db	GCCCGATCTTCTAAACAACTTTTGGGAAGTTATTGAGGTGAAGCCGACCAAGAGAACC	1320
1321	Qy	CCAGCTTCAGGACCTGGACATGATGCTCACTCGTGTAGAGATCAAGAAAGGAGCTTTGAT	1380
1321	Db	CCAGCTTCAGGACCTGGACATGATGCTCACTCGTGTAGAGATCAAGAAAGGAGCTTTGAT	1380

Qy	1381	GCTCCACACTTCAA	CTCAAAGGCCATGGTTATCGTCGTCTCGTCAAAGGAACTGAAA	1440
Db	1381	GCTCCACACTTCAA	CTCAAAGGCCATGGTTATCGTCGTCTCGTCAAAGGAACTGAAA	1440
Qy	1441	CCTTGAACCTCGTGG	CTGTTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAAAGAGAGGA	1500
Db	1441	CCTTGAACCTCGTGG	CTGTTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAAAGAGAGGA	1500
Qy	1501	GGACGAGACGAGAG	AGAGGAGGAGTAACACAGAGAGGTGCGTAGGTACACACGCGAGTT	1560
Db	1501	GGACGAGACGAGAG	AGAGGAGGAGTAACACAGAGAGGTGCGTAGGTACACACGCGAGTT	1560
Qy	1561	GAAGGAAGGCGATG	TGTTCAATCGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Db	1561	GAAGGAAGGCGATG	TGTTCAATCGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Qy	1621	CGAACTCCATCTG	CTTTGGCTTCGGTATCAACCGCTGAAAAACAACACAGAAATCTTCCTTGC	1680
Db	1621	CGAACTCCATCTG	CTTTGGCTTCGGTATCAACCGCTGAAAAACAACACAGAAATCTTCCTTGC	1680
Qy	1681	AGTGATTAAGCAAT	TGTGATAGACCAAGATAGAGACGAGCAAGCAAGGATTTAGCATTCCC	1740
Db	1681	AGTGATTAAGCAAT	TGTGATAGACCAAGATAGAGACGAGCAAGCAAGGATTTAGCATTCCC	1740
Qy	1741	TGGGTGGGTGAACA	AGTTTGAGAAAGTCTCAATCAAAAAACGAGAGGAATCTCACATTTGTGAG	1800
Db	1741	TGGGTGGGTGAACA	AGTTTGAGAAAGTCTCAATCAAAAAACGAGAGGAATCTCACATTTGTGAG	1800
Qy	1801	TGTCGTCTCCTCA	ATCTCAATCTCAATCTCGTCGTCTCCTGAGAAAAGAGTCTCCTGAGAA	1860
Db	1801	TGTCGTCTCCTCA	ATCTCAATCTCAATCTCGTCGTCTCCTGAGAAAAGAGTCTCCTGAGAA	1860
Qy	1861	AGAGGATCAAGAG	GAGAAAAACAAGAGGGAAGGTCCACTCTTTCAATTTTGAAGGC	1920
Db	1861	AGAGGATCAAGAG	GAGAAAAACAAGAGGGAAGGTCCACTCTTTCAATTTTGAAGGC	1920
Qy	1921	TTTTTAAC	TGA 1930	
Db	1921	TTTTTAAC	TGA 1930	
RESULT 7				
AR257463				
LOCUS	AR257463	1930 bp	DNA	linear PAT 20-DEC-2002
DEFINITION	Sequence 3 from patent US 6486311.			
ACCESSION	AR257463			
VERSION	AR257463.1	GI:27307476		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1930)			
	Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,			
	Helm,R.M. and Bannon,G.A.			
TITLE	Patent allergens and methods			
JOURNAL	Patent: US 6486311-A 3 26-NOV-2002;			
FEATURES	Location/Qualifiers			
source	1..1930			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Query Match	95.0%;	Score 1930;	DB 6;	Length 1930;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1930;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	AATAATCATATATTTCATCATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Db	1	AATAATCATATATTTCATCATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Qy	61	GGTTTCTCCACTGATGCTGTTGCTAGGAGTCCCTTGTCTCGGCTTCAGTTTCTGCAACGCA	120	

Db	61	GGTTTCTCCACATGATGCTGTTGTCTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCCA	120
Qy	121	TGCCAAGTCATCACCTTTACCAAGAGAAGAAAACAGAGAAACCCCTGCGCCACAGAGGTGCCTCCA	180
Db	121	TGCCAAGTCATCACCTTTACCAAGAGAAGAAAACAGAGAAACCCCTGCGCCACAGAGGTGCCTCCA	180
Qy	181	GAGTTGTCAACAGGAAACCGGATGACTTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240
Db	181	GAGTTGTCAACAGGAAACCGGATGACTTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240
Qy	241	GCTCGAGTATGATCCTCGTTGTGTCATGATCTCTCGAGGACACACTCGCAACCAACCA	300
Db	241	GCTCGAGTATGATCCTCGTTGTGTCATGATCTCTCGAGGACACACTCGCAACCAACCA	300
Qy	301	ACGTTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCGGAGACTTACGATGATGACCG	360
Db	301	ACGTTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCGGAGACTTACGATGATGACCG	360
Qy	361	CCGTCAACCCCGAAGAGAGGAGCGCCGATGCGGACCACTGCGAACCGAGGGAGCGTGA	420
Db	361	CCGTCAACCCCGAAGAGAGGAGCGCGATGCGGACCACTGCGAACCGAGGGAGCGTGA	420
Qy	421	AAGAGAGAGACTTGGAGACAACCAAGAGAGATTGGAGCGGACCAAGTCATCAGACGCC	480
Db	421	AAGAGAGAGACTTGGAGACAACCAAGAGAGATTGGAGCGGACCAAGTCATCAGACGCC	480
Qy	481	ACGGAATAAGGCCGCAAGGAAGAGAGAAACAAGATGGGGAAACACGAGTAGGCCA	540
Db	481	ACGGAATAAGGCCGCAAGGAAGAGAGAAACAAGATGGGGAAACACGAGTAGGCCA	540
Qy	541	TGTGAGGGAAGAAAACATCTCGGAAACAACCTTTCTACTTCCGTCAAAGCGGTTTAGCAC	600
Db	541	TGTGAGGGAAGAAAACATCTCGGAAACAACCTTTCTACTTCCGTCAAAGCGGTTTAGCAC	600
Qy	601	CCGCTACGGGAACAAAACGCTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTTCAAG	660
Db	601	CCGCTACGGGAACAAAACGCTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTTCAAG	660
Qy	661	GCAGTTTCAGAACTCTCCAGAACTACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCT	720
Db	661	GCAGTTTCAGAACTCTCCAGAACTACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCT	720
Qy	721	TGTTCTTTCCCAAGCACGCTGATGATAACATCCTTTGTTATCAGCAAGGGCAAGCCAC	780
Db	721	TGTTCTTTCCCAAGCACGCTGATGATAACATCCTTTGTTATCAGCAAGGGCAAGCCAC	780
Qy	781	CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAACTTGTGACAGGGCCATGCACT	840
Db	781	CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAACTTGTGACAGGGCCATGCACT	840
Qy	841	CAGAACTCCATCCGTTTTCATTTTCTTACATCTTCAACCGCCATCACAACAGAACCTTCAG	900
Db	841	CAGAACTCCATCCGTTTTCATTTTCTTACATCTTCAACCGCCATCACAACAGAACCTTCAG	900
Qy	901	AGTAGCTAAATCTCCATGCGCGTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC	960
Db	901	AGTAGCTAAATCTCCATGCGCGTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC	960
Qy	961	GAGCAGCGGAGACCAATCATCTACTTGTGAGGGCTTCAGCAGGAATACGTTGGAGGCGCG	1020
Db	961	GAGCAGCGGAGACCAATCATCTACTTGTGAGGGCTTCAGCAGGAATACGTTGGAGGCGCG	1020
Qy	1021	CTTCAATGCGGAATTCAATGAGATACGAGGGGTCTGTTAGAAGAGAAATGCAGGAGTGA	1080
Db	1021	CTTCAATGCGGAATTCAATGAGATACGAGGGGTCTGTTAGAAGAGAAATGCAGGAGTGA	1080
Qy	1081	GCAAGAGGAGAGGCGCAGGCGCATCGGAGTACTCGGAGTAGTCAGAAACAATGAAGGAGT	1140
Db	1081	GCAAGAGGAGAGGCGCAGGCGCATCGGAGTACTCGGAGTAGTCAGAAACAATGAAGGAGT	1140
Qy	1141	GATAGTCAAGTGTCAAAGGAGCAAGTTTAAGAACTTACTTAAGCAACGCTTAATCCGCTCTC	1200
Db	1141	GATAGTCAAGTGTCAAAGGAGCAAGTTTAAGAACTTACTTAAGCAACGCTTAATCCGCTCTC	1200

QY	1201	AAAGAAAGGCTCCGAAGAAGAGGAGGATATCA	CCAAACCCAAATCAACTTGTAGAGAAGCGGA	1260
DB	1201	AAAGAAAGGCTCCGAAGAAGAGGAGGATATCA	CCAAACCCAAATCAACTTGTAGAGAAGCGGA	1260
QY	1261	GCCCGATCTTTCTAAACAACCTTTGGGAAGTTAT	TTTGTAGGTGAAGCCAGACAAGAGAAACCC	1320
DB	1261	GCCCGATCTTTCTAAACAACCTTTGGGAAGTTAT	TTTGTAGGTGAAGCCAGACAAGAGAAACCC	1320
QY	1321	CCAGCTTTCAGGACCTGGACATGATGCTTCAC	CTGTGTAGAGATCAAGAAGGAGCTTTGTAT	1380
DB	1321	CCAGCTTTCAGGACCTGGACATGATGCTTCAC	CTGTGTAGAGATCAAGAAGGAGCTTTGTAT	1380
QY	1381	GCTCCCACTTCAACTCAAAAGGCCATGGTTAT	TCGTCTGTCTCAACAAAGGAACTGGAAA	1440
DB	1381	GCTCCCACTTCAACTCAAAAGGCCATGGTTAT	TCGTCTGTCTCAACAAAGGAACTGGAAA	1440
QY	1441	CTTTGAACCTCGTGCTCTAAGAAAGAGCAACA	CACAGAGGGGACGCGGGAAGAAGAGGA	1500
DB	1441	CTTTGAACCTCGTGCTCTAAGAAAGAGCAACA	CACAGAGGGGACGCGGGAAGAAGAGGA	1500
QY	1501	GGACGAAGAAGCAAGAAGAGGAGGGAAGTAA	CACAGAGGTGCGTAGGTACACAGCGAGGTT	1560
DB	1501	GGACGAAGAAGCAAGAAGAGGAGGGAAGTAA	CACAGAGGTGCGTAGGTACACAGCGAGGTT	1560
QY	1561	GAAGGAAGGCGATGTGTTTCATCATGCGACAG	CTCATCAGTAGGCCATCAACGCTTCCCTC	1620
DB	1561	GAAGGAAGGCGATGTGTTTCATCATGCGACAG	CTCATCAGTAGGCCATCAACGCTTCCCTC	1620
QY	1621	CGAACTCCATCTCGCTTCGGCTTTCGGTATCA	ACGCTGAAACAACCAAGAAATCTTCTCTGC	1680
DB	1621	CGAACTCCATCTCGCTTCGGCTTTCGGTATCA	ACGCTGAAACAACCAAGAAATCTTCTCTGC	1680
QY	1681	AGGTGATAGGACAATGTGATAGACCAAGTAGA	AGCAAGCAAGCAAGGATTTAGCATTTCCC	1740
DB	1681	AGGTGATAGGACAATGTGATAGACCAAGTAGA	AGCAAGCAAGCAAGGATTTAGCATTTCCC	1740
QY	1741	TGGGTCCGGTGAAACAGTTGAGAAAGCTCAT	CAAAAACCAAGAGGAATCTCATTGTGTAG	1800
DB	1741	TGGGTCCGGTGAAACAGTTGAGAAAGCTCAT	CAAAAACCAAGAGGAATCTCATTGTGTAG	1800
QY	1801	TGCTCGTCCCTCAATCTCAATCTCAATCTCGT	CTCGTCTCCTGTAGAAAGAGTCTCCTGTAGAA	1860
DB	1801	TGCTCGTCCCTCAATCTCAATCTCAATCTCGT	CTCGTCTCCTGTAGAAAGAGTCTCCTGTAGAA	1860
QY	1861	AGAGGATCAAGAGGAGGAAAAACCAAGGAGGA	AGGGTCCACTCCTTTCAATTTTGAAGGC	1920
DB	1861	AGAGGATCAAGAGGAGGAAAAACCAAGGAGGA	AGGGTCCACTCCTTTCAATTTTGAAGGC	1920
QY	1921	TTTTTAACCTGA	1930	
DB	1921	TTTTTAACCTGA	1930	

RESULT 8
RD107898

BD107898 1930 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION
Methods and reagents for decreasing allergic reactions.
BU167636

DEFINITION METHODS and references to
ACCESSION BP107898

ACCESSION	BD107898
VERSION	BD107898.1
	GT:23202716

VERSION
KEYWORDS

KEYWORDS
SOURCE
OF 2002501740
unidentified

SOURCE	unidentified
ORGANISM	unidentified

ORGANISM unidentified
 unclassified

REFERENCE 1 (pages 1 to 1930)
UNCLASSIFIED.

REFERENCE
I (bases I to 1930)
Cosin H, Benson C A, Jr W A B and Samanthon H A
MITHOPS

AUTHORS	TITLE
Sobin, H., Baron, G.A., Jr, W.A.B. and Sal...	Methods and concepts for detecting al...

TITLE Methods and reagents for decreasing allergic reactions

JOURNAL
Patent: JP 2002501748-A 1 22-JAN-2003
INSTITUTE OF ADVANCED SCIENCE AND TECHNOLOGY

UNIVERSITY OF ARKANSAS,
FACULTY OF NEW YORK

UNIVERSITY OF NEW YORK

COMMENT	OS	Arachis L.

[illegible]

QY	1561	GAAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Db	1535	GAAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC	1594
QY	1621	CGAATCTCMTCTGCTTGGCTTCGGTATCAACCTGAAACAAACAGAAATCTTCCTTGC	1680
Db	1595	CGAATCTCMTCTGCTTGGCTTCGGTATCAACCTGAAACAAACAGAAATCTTCCTTGC	1654
QY	1681	AGTGATAGGACAAATGTATAGACAGATAGAGCAAGAGCAAGCAAGGATTTAGCAATCCC	1740
Db	1655	AGTGATAGGACAAATGTATAGACAGATAGAGCAAGCAAGGATTTAGCAATCCC	1714
QY	1741	TGGGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAG	1800
Db	1715	TGGTTCGGGTGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAG	1774
QY	1801	TGCTGCTCCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAAGAGATCTCTTGAGAA	1860
Db	1775	TGCTGCTCCTCAATCTC-----CGTGTCTCTCTGAATA	1806
QY	1861	AGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC	1920
Db	1807	AGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC	1866
QY	1921	TTTTAACTGAGATGGAGCAACTTGTATGTATCGATAATAAGATCACGCTTTTGTACT	1980
Db	1867	TTTTAACTGAGATGGAGCAACTTGTATGTATCGATAATAAGATCACGCTTTTGTAAAT	1926
QY	1981	CTACTATCCAAAACCTTATCAATAATAATAAAACGTTTGTGGTGTCTCTCC	2032
Db	1927	CTACTATCCAAAACCTTATCAATAATAATAAAACGTTTGTGGTGTCTCTCC	1978
RESULT 10			
AR257467			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1994; Conservative			
QY	48	CAATGAGAGGAGGGTTCCTCCACGTGATGCTGTTCCTAGGATCCTGTCTGGCTTCAG	107
Db	1	CAATGAGAGGAGGGTTCCTCCACGTGATGCTGTTCCTGGATCCTGTCTGGCTTCAG	60
QY	108	TTTCTGCAACGCATCCCAAGTCATCACCTTACCAGAAAGAACAGAGAACCCCTCGGCC	167
Db	61	TTTCTGCAACGCATCCCAAGTCATCACCTTACCAGAAAGAACAGAGAACCCCTCGGCC	114
QY	168	AGAGTGTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCCGAGT	227
Db	115	AGAGTGTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCCGAGT	174
QY	228	CTCGTGCACCAAGCTCAGATATGATCCTCGTTGTGTCTATGATCCTCGAGACACACTG	287
Db	175	CTCGTGCACCAAGCTCAGATATGATCCTCGTTGTGTCTATG-----ACACTG	222

QY	288	GCACCACCAACCAACGTTCCCTTCAGGGGAGCGACACGTGGCGCCAAACCCGAGACT	347
Db	223	GGCCACCAACCAACGTCACCTTCAGGGGAGCGGACACGTGGCGCCAAACCCGAGACT	282
QY	348	ACGATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGGAACAGCTGGAC	407
Db	283	ACGATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGGAACAGCTGAAC	342
QY	408	CGGGGAGCGTGAAGAGAAAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCA	467
Db	343	CGGGGAGCGTGAAGAGAAAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCA	402
QY	468	GTCAATCAGCACCGACCGAAAAATAAGGCCCGAAGAGAGAGGAGAAACAAGAGTGGGAA	527
Db	403	GTCAATCAGCACCGACCGAAAAATAAGGCCCGAAGAGAGAGGAGAAACAAGAGTGGGAA	462
QY	528	CACAGGTAGCATGTGTAGGGAAGAAACATCTCGGAAACAACCTTTTACTTTCCCGTCAA	587
Db	463	CACAGGTAGCATGTGTAGGGAAGAAACATCTCGGAAACAACCTTTTACTTTCCCGTCAA	522
QY	588	GGCGTTTAGCACCCGCTACCGGAAACCAAAACGCTAGGATCGGGTCTCGCAGAGTTTG	647
Db	523	GGCGTTTAGCACCCGCTACCGGAAACCAAAACGCTAGGATCGGGTCTCGCAGAGTTTG	582
QY	648	ACCAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCA	707
Db	593	ACCAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCA	642
QY	708	AACCTAACACACTCTTTGTTCTTCCCAAGCACGCTGATGCTGATAACAATCCTTGTATCCAGC	767
Db	643	GACCTAACACACTCTTTGTTCTTCCCAAGCACGCTGATGCTGATAACAATCCTTGTATCCAGC	702
QY	768	AAGGCAAGCCACCGTAGCCGTAAGTAATGCAATTAACAGAAAGAGCTTTAATCTTGACG	827
Db	703	AAGGCAAGCCACCGTAGCCGTAAGTAATGCAATTAACAGAAAGAGCTTTAATCTTGACG	762
QY	828	AGGCGATGCACTCAGAAATCCATCCGTTTCATTCTCTACATCTTGAACCGCATGACA	887
Db	763	AGGCGATGCACTCAGAAATCCATCCGTTTCATTCTCTACATCTTGAATCGACATGACA	822
QY	888	ACCAGAACCTCAGAGTAGCTAAATTCATTCATCCCGTTTAAACACACCGCGCAGTTTGAGG	947
Db	823	ACCAGAACCTCAGAGTAGCTAAATTCATTCATCCCGTTTAAACACACCGCGCAGTTTGAGG	882
QY	948	ATTTCTTCCCGCGGAGCGAGCCAGACCAATCATCTCTACTTCAGGGCTTCAGCAGGATA	1007
Db	883	ATTTCTTCCCGCGGAGCGAGCCAGACCAATCATCTCTACTTCAGGGGATTCAGCAGGATA	942
QY	1008	CGTTGGAGCGCCCTTCAATCGGAAATTCATATGAGATACGAGGGTGTCTGTTAGAAGAGA	1067
Db	943	CGTTGGAGCGCCCTTCAATCGGAAATTCATATGAGATACGAGGGTGTCTGTTAGAAGAGA	1002
QY	1068	ATGAGGAGGTGAGCAAGAGAGAGAGGCGAGCGATGAGTACTCGGAGTAGTGAGA	1127
Db	1003	ATGAGGAGGTGAGCAAGAGAGAGAGGCGAGCGAGCGAGTACTCGGAGTAGTGAGA	1059
QY	1128	ACAATGAAGAGTATAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTAAGCAGC	1187
Db	1060	ATAATGAAGAGTATAGTCAAGTGTCAAGAGGACGCTTCAAGAACTTACTAAGCAGC	1119
QY	1188	CTAATTCGCTCTCAAGAAAGGCTCCGAAAGAGGAGAGATATCAACCAACCAATCAACT	1247
Db	1120	CTAATTCGCTCTCAAGAAAGGCTCC---GAAAGAGAGATATCAACCAACCAATCAACT	1176
QY	1248	TGAGAGAGGCGAGCCCGATCTTTTAAACAATTTGGGAACTTATTTGAGGTGAAGCAG	1307
Db	1177	TGAGAGATGGGAGCCCGATCTTTTAAACAATTTGGGAGTTTATTTGAGGTGAAGCAGC	1236
QY	1308	ACAAGAGAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGCTAGAGATCAAG	1367
Db	1237	ACAAGAGAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGCTAGAGATCAAG	1296

```
QY 1368 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACA 1427
Db 1297 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACA 1356
QY 1428 AAGGAACTGGAACCTTTGAACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGACGGC 1487
Db 1357 AAGGAACTGGAACCTTTGAACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGACGGC 1416
QY 1488 GGGAA-----GAAGAGGAGCAGACGACGAGAGGAGGAGGAGGAGGAGGAGGAGG 1538
Db 1417 GGGAAACAAGAGTGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1476
QY 1539 TGGGTAGGTACACAGGAGGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1598
Db 1477 TGGGTAGGTACACAGGAGGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1536
QY 1599 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGGTATCAACGCTGAAA 1596
QY 1659 ACAACACAGAAATCTCTTGGAGGTGATAGGACAAATGTGATAGACCAAGATAGAGAAGC 1718
Db 1597 ACAACACAGAAATCTCTTGGAGGTGATAGGACAAATGTGATAGACCAAGATAGAGAAGC 1656
QY 1719 AAGCGAAGGATTTAGCATTCCTCGGCTCGGTCGAGTCAAGTGAAGAGCTCATCAAAAACC 1778
Db 1657 AAGCGAAGGATTTAGCATTCCTCGGCTCGGTCGAGTCAAGTGAAGAGCTCATCAAAAACC 1716
QY 1779 AGAAGGAATCTCACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCCGCTCTC 1838
Db 1717 AGAGGAGTCTCACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCCGCTC----- 1765
QY 1839 CTGAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1898
Db 1766 -----GTCTCTGAAAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1815
QY 1899 CACTCTCTCAATTTGAGGCTTTAACTGAGAAATGGAGGCAACTGTTATGTATCGAT 1958
Db 1816 CACTCTCTCAATTTGAGGCTTTAACTGAGAAATGGAGGCAACTGTTATGTATCGAT 1875
QY 1959 AATAAGATCAGCTTTGTACTCTACTATCAAAAACTTATCAATAAATAAAAACTTTG 2018
Db 1876 AATAAGATCAGCTTTGTATCTACTATCAAAAACTTATCAATAAATAAAAACTTTG 1935
QY 2019 TGGTTTGTCTCC 2032
Db 1936 TGGTTTGTCTCC 1949

RESULT 11
ARQARAH
LOCUS
DEFINITION
Arachis hypogaea (clone P17) Ara h I mRNA, linear, PLN 10-JAN-1995
ACCESSION
L38853
VERSION
L38853.1 GI:620024
KEYWORDS
peanut hypersensitivity.
SOURCE
Arachis hypogaea
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 1949)
Burks, A.W., Cockrell, G., Stanley, J.S., Helm, R.M. and Bammon, G.A.
Recombinant peanut allergen Ara h I expression and IgE binding in
patients with peanut hypersensitivity
JOURNAL
Unpublished (1994)
COMMENT
Original source text: Arachis hypogaea (strain Florunner) Seed cDNA
to mRNA.
FEATURES
Location/Qualifiers
1..1949
/organism="Arachis hypogaea"
/mol_type="mRNA"
```

```
/strain="Florunner"
/db_xref="taxon:3818"
/dev_stage="Seed"
1..1949
/gene="Ara h I"
1..2
/gene="Ara h I"
3..1847
/gene="Ara h I"
/codon_start=1
/protein_id="AAA60336.1"
/db_xref="GI:620025"
/translation="MRGRVSPLMLLGLILVASVATQAKSPYRKTEPCQRCLOSC
QEPDLDKQKACESRCTKLEYDPRCYDGTGNQRHPGERTGRQPGDYDDRRRQPR
REGGRWGPAPREPEREDWRQPRDWRPQPRKIPRPGREGQEGWGTGSEVR
ESTSNPNFPFSTRYNGONGRIRVLQRFQDQSKQFQNLQNHRIYQIENRPNL
VLPIKADADNLIQOQATVTVAANGNNKSFNLDEGHALRIPSGTISVILNRHDNQ
LRVAKISMPVNTPGQEDFPFASRDQSYLOGFSRNTLEAFAFNRIIRVLRLEN
AGGEQERQORERTSRSDNEGVIKVSKHEVQELTKHAKSVSKGSESEDITNPLN
RGEEDLNNFGRLEFVPRDKNPQLQDLMLTCEIKEGALMLPHFNSKAMVIVV
NKGTLNLELVAVRQKQRRQREQBEDEEBEGSNREVRRTARLKGDDVFTMP
AAHPVAINASSELHLLGFINAENNRIFLAGDKONVIDOIIEKQADLAFPGSGEQVE
KLIKQNRSHFVSARPOSQSPSPPEKEDQEEENQGGKGPLLSILKAFN"
1848..1949
3' UTR
/gene="Ara h I"
1949
polyA_site
/gene="Ara h I"
1949
ORIGIN
Query Match 86.2%; Score 1752.4; DB 8; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;
QY 48 CAATGAGAGGAGGGTTTCTCCACATGATCTGTGTCTAGGGATCTTGTCTGGCTTCAG 107
Db 1 CAATGAGAGGAGGGTTTCTCCACATGATCTGTGTCTGGATCTTGTCTGGCTTCAG 60
QY 108 TTTCTGCAACGATGCGCAAGTCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167
Db 61 TTTCTGCAACGAGGCGCAAGTCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 114
QY 168 AGAGGTGCTCCAGAGTGTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
Db 115 AGAGGTGCTCCAGAGTGTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174
QY 228 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTGTATGATCTCTCGAGGACACATG 287
Db 175 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTGTATG-----ACACTG 222
QY 288 GCACCAACCAACGTTTCCCTCCAGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 347
Db 223 GCGCCACCAACCAACGTTTCCCTCCAGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 282
QY 348 AGGATGATGACCGCGCTCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407
Db 283 AGGATGATGACCGCGCTCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 342
QY 408 CGAGGAGAGCGGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
Db 343 CGAGGAGAGCGGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
QY 468 GTTCATCAGAGCCACCGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 527
Db 403 GTTCATCAGAGCCACCGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
QY 528 CACCAAGTATGAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 587
Db 463 CACCAAGTATGAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
QY 588 GCGCGTTTACGACCGCTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 647
Db 523 GCGCGTTTACGACCGCTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
```


exon	883. .1058	KAMVIVVVKGTGNLELVAVRKQQQRGRREDEDEEEEGSNREVRVRYTARKEGD
	/number=2	
exon	1308. .1388	VFTMPAHPVAIIINASSLHLGFGINAENHRIFLAGDKONVIDQIEKQAKDLAFPGS
	/number=3	GEQVEKLIIKNQKESHFVSAR PQSQSPSPSEKESPEKED EENQGGKPLLSILKA
exon	1463. .2674	FN"
	/number=4	
3'UTR	2366. .2674	
polyA_signal	2440. .2445	
ORIGIN		

Query Match	79.1%;	Score 1608;	DB 8;	Length 2674;
Best Local Similarity	83.8%;	Pred. No. 0;		
Matches 2012;	Conservative 0;	Mismatches 0;	Indels 394;	Gaps 3;
Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGTAGCAGGAGCAATGAGAGGGAG	60	
Db	42	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGTAGCAGGAGCAATGAGAGGGAG	101	
Qy	61	GGTTTCTCCACTGATGCTGTTGCTTAGGGATCCTTGTCTCGCTTCAGTTCTTCGCAACGCA	120	
Db	102	GGTTTCTCCACTGATGCTGTTGCTTAGGGATCCTTGTCTCGCTTCAGTTCTTCGCAACGCA	161	
Qy	121	TGCCAAGTCATCACTTTACAGAGAAGAAAAAGAGAACCCCTGCGCCAGAGGTGCCTCCCA	180	
Db	162	TGCCAAGTCATCACTTTACAGAGAAGAAAAAGAGAACCCCTGCGCCAGAGGTGCCTCCCA	221	
Qy	181	GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAA	240	
Db	222	GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAA	281	
Qy	241	GCTCGAGTATGATCCTCGTTGTGTCATGATCCTCGAGGACACACTGCGACCAACCAACCA	300	
Db	282	GCTCGAGTATGATCCTCGTTGTGTCATGATCCTCGAGGACACACTGCGACCAACCAACCA	341	
Qy	301	ACGTTCCCTCCAGGGAGCGGACACGTGCGCGCCAAACCGCGAGACTACGATGATGACCG	360	
Db	342	ACGTTCCCTCCAGGGAGCGGACACGTGCGCGCCAAACCGCGAGACTACGATGATGACCG	401	
Qy	361	CCGTCAACCCCGAAGAGAGGAAGAGGCCGATGGGGAACGAGTTGACCGAGGAGCGTGA	420	
Db	402	CCGTCAACCCCGAAGAGAGGAAGAGGCCGATGGGGAACGAGTTGACCGAGGAGCGTGA	461	
Qy	421	AAGAGAGAGAGACTGGAGACAAACAGAGAAAGATTGGAGCGGACCAAGTCATCAGCAGCC	480	
Db	462	AAGAGAGAGAGACTGGAGACAAACAGAGAAAGATTGGAGCGGACCAAGTCATCAGCAGCC	521	
Qy	481	ACGGAAAAATAGGCCCGAAGAGAGAGAGAAAGAGTGGGGAAACACAGGTAGCCCA	540	
Db	522	ACGGAAAAATAGGCCCGAAGAGAGAGAGAGAAAGAGTGGGGAAACACAGGTAGCCCA	581	
Qy	541	TGTTAGGGAAAGAAACATCTCGGAAACAAACCTTTCTACTTCCGTCGAGCGGTTTAGCAC	600	
Db	582	TGTTAGGGAAAGAAACATCTCGGAAACAAACCTTTCTACTTCCGTCGAGCGGTTTAGCAC	641	
Qy	601	CCGCTACGGGAAACCAAAACGGTAGGATCCGGGTCTCTGACAGGTTTGACCAAAAGGTCAG	660	
Db	642	CCGCTACGGGAAACCAAAACGGTAGGATCCGGGTCTCTGACAGGTTTGACCAAAAGGTCAG	701	
Qy	661	GCAGTTTCAGAACTCCAGAAATCACCGTATTTGTGCGAGATCGAGGCCAAACCTAACACTCT	720	
Db	702	GCAGTTTCAGAACTCCAGAAATCACCGTATTTGTGCGAGATCGAGGCCAAACCTAACACTCT	761	
Qy	721	TGTTCTTCCCAAGCAGCGCTGATGCTGATAACATCCTTTGTTATCCAGCA	768	
Db	762	TGTTCTTCCCAAGCAGCGCTGATGCTGATAACATCCTTTGTTATCCAGCAAGGATCAAAATC	821	
Qy	769	-----A	769	
Db	822	TAATTCCTATTCAAACCTACATATATTTGTTGCTTGATACATATGATTCATTGGATTGCA	881	

QY	770	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAAACAGAAAGAGCTTTAAATCTTTGACGAG	829
DB	882	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAAACAGAAAGAGCTTTAAATCTTTGACGAG	941
QY	830	GGCCATGCACCTCAGAAATCCCAATCCGGTTTTCAATTTCTTACATCTTTGAAACCGGCATGACAAAC	889
DB	942	GGCCATGCACCTCAGAAATCCCAATCCGGTTTTCAATTTCTTACATCTTTGAAACCGGCATGACAAAC	1001
QY	890	CAGAACTCTCAGAGTAGCTTAAAAATCTCCATGCCCGTTTAAACACACCCGGCCAGTTTG-	944
DB	1002	CAGAACTCTCAGAGTAGCTTAAAAATCTCCATGCCCGTTTAAACACACCCGGCCAGTTTGAGGTA	1061
QY	945	-----	944
DB	1062	CCTCTTCTTTTACATATATATTTCAAGAGGAATGTTAGGAGGCCAGCACTTTTGTAAAT	1121
QY	945	-----	944
DB	1122	TTGTAGTTATCAAAATAGTCATTAGTGATGATATTAAATGTTATTAAATAGATTTTAAATGG	1181
QY	945	-----	944
DB	1182	TGTAAATTTTATCTAATGGTTCAATTTTTTTTTTTTGGATACATGTTGGTCAGAAATTTAAAC	1241
QY	945	-----	944
DB	1242	AAATTACTGGTCTAGATTTTTTCTTATATTCAATTTCTCAATTAATCATCTTTACATGTTGCT	1301
QY	945	----AGAAATTTCTCCGGGAGCAGCGGAGACCAATCATCTACTTTGCGAGGGCTTCAGC	1000
DB	1302	TCACAGGATTTCTTTCCGGGAGCAGCGGAGACCAATCATCTACTTTGCGAGGGCTTCAGC	1361
QY	1001	AGGAATACGTTGGAGGCCGCTTCAAT-----	1027
DB	1362	AGGAATACGTTGGAGGCCGCTTCAATGTAAGCAAAATGTATCATAAATTAGGAATTTAAAT	1421
QY	1028	-----GCGGAATTTCAATGAGATAC	1046
DB	1422	AACGTATATCTGTTATAATAAACTTATTAATATATACATAGGCGGAATTTCAATGAGATAC	1481
QY	1047	GGAGGTCGTTTGAAGAGNAATGCAGGAGGTGACGAGGAGGAGAGGCGCAGAGGCGAT	1106
DB	1482	GGAGGTCGTTTGAAGAGNAATGCAGGAGGTGACGAGGAGGAGAGGCGCAGAGGCGAT	1541
QY	1107	GGAGTACTCGGAGTAGTGAGAAACAAATCAAGGAGGTAGTAGTCAAAAGTGTCAAAGGAGACAGC	1166
DB	1542	GGAGTACTCGGAGTAGTGAGNAACAAATCAAGGAGGTAGTAGTCAAAAGTGTCAAAGGAGACAGC	1601
QY	1167	TTGAAGAACTTTAATAAGCAACGTTAAATCCGTTCTCAAAAGAAAGGCTCCGAAGAAGAGGGAG	1226
DB	1602	TTGAAGAACTTTAATAAGCAACGTTAAATCCGTTCTCAAAAGAAAGGCTCCGAAGAAGAGGGAG	1661
QY	1227	ATATCACCAACCCCAATCAACTTGAGAGAGGCGAGCCGATCTTTCTTCAAACACTTTCCGGA	1286
DB	1662	ATATCACCAACCCCAATCAACTTGAGAGAGGCGAGCCGATCTTTCTTCAAACACTTTCCGGA	1721
QY	1287	AGTTATTTTGGGTGAAGCCAGACAAGAAAGAACCCCGAGTTTCAGGACTCGGACATGATGC	1346
DB	1722	AGTTATTTTGGGTGAAGCCAGACAAGAAAGAAACCCCGAGTTTCAGGACTCGGACATGATGC	1781
QY	1347	TCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCCACTTCCACACTTCAAAGGCCA	1406
DB	1782	TCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCCACTTCCACACTTCAAAGGCCA	1841
QY	1407	TGCTTATCGTCGTCTCAAACAAAGGAACCTTGAAACCTTGAACTCGTGGCTGTGAAGAAAG	1466
DB	1842	TGCTTATCGTCGTCTCAAACAAAGGAACCTTGAAACCTTGAACTCGTGGCTGTGAAGAAAG	1901
QY	1467	AGCAACAAACAGAGGGGACCGCGGGAGAGAGAGGAGGACGAAGACGAAGAGAGAGGGAA	1526
DB	1902	AGCAACAAACAGAGGGGACCGCGGGAGAGAGAGGAGGACGAAGACGAAGAGAGAGGGAA	1961

QY 1527 GTAACAGAGAGGTGGCTAGGTACACAGCAGAGGTGTAAGGAAGCGATGTGTTTCATCATGC 1586
Db GTAAACAGAGAGGTGGCTAGGTACACAGCAGAGGTGTAAGGAAGCGATGTGTTTCATCATGC 2021
QY 1587 CAGCAGCTCATCAGTAGGACATCAAGCGTTCTCCGAACTCATCTGCTTGGCTTCGGTA 1646
Db CAGCAGCTCATCAGTAGGACATCAAGCGTTCTCCGAACTCATCTGCTTGGCTTCGGTA 2081
QY 1647 TCAAGCGCTCAAAACACACAGAACTCTCTTGCAGGTGATAGGACAAATGATAGACC 1706
Db TCAAGCGCTCAAAACACACAGAACTCTCTTGCAGGTGATAGGACAAATGATAGACC 2141
QY 1707 AGATAGAGAGCAAGCAAGAGATTTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAAGC 1766
Db AGATAGAGAGCAAGCAAGAGATTTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAAGC 2201
QY 1767 TCATCAAAACACAGAAAGAACTCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAAT 1826
Db TCATCAAAACACAGAAAGAACTCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAAT 2261
QY 1827 CTCCTGCTCTCTGAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAG 1886
Db CTCCTGCTCTCTGAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAG 2321
QY 1887 GAGGAAGGCTCACTCTCTTCAATTTTGAAGGCTTTTAACTGAGAAAGGAGGCAACTTG 1946
Db GAGGAAGGCTCACTCTCTTCAATTTTGAAGGCTTTTAACTGAGAAAGGAGGCAACTTG 2381
QY 1947 TTATGATCGAATAAGATACGCTTTTGTACTCTACTATCAAAACTTATCAATAA 2006
Db TTATGATCGAATAAGATACGCTTTTGTACTCTACTATCAAAACTTATCAATAA 2441
QY 2007 TAAAAACGTTTGTGCGTGTCTTCTCC 2032
Db TAAAAACGTTTGTGCGTGTCTTCTCC 2467

RESULT 13

AY581852
LOCUS AY581852 2332 bp DNA linear PLN 01-MAY-2004
DEFINITION Arachis hypogaea conarachin gene, partial cds.
ACCESSION AY581852
VERSION AY581852.1 GI:46560477
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE 1 (bases 1 to 2332)
AUTHORS Li,H., Wang,L., Liao,B., Yan,Y., Lin,X. and Huang,S.
TITLE Molecular cloning of conarachin gene from peanut seed

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2332)
AUTHORS Li,H., Wang,L., Liao,B., Yan,Y., Lin,X. and Huang,S.
TITLE Direct Submission

JOURNAL Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
GuangDong 510275, China
Location/Qualifiers

FEATURES
source

1..2332
/organism="Arachis hypogaea"
/mol_type="genomic DNA"
/cultivar="Sanfou523"
Join(343..1045,1117..1292,1369..1449,1523..>2303)
/product="conarachin"
Join(343..1045,1117..1292,1369..1449,1523..>2303)
/codon_start=1
/product="conarachin"
/protein_id="AA00597.1"
/db_xref="GI:46560478"

mRNA

CDS

/translation="MRGRVSPIMLLILVLASVATQAKSPYRKTNPCQAQCQSC
QOEPPDLKOKACSRCTKLEYPDPCVDTGATNQHPGERTRQPGQDYDDRRQPR
REEGMDGAPREREREEDWOPREDWRRDSHOQPKIRPEGREGEWEGWTPGSEVR
BETSRNPPFPSPRSTRYCNQNGRIRVLQRPQORSQFQNLQNRHIVQLEARNPIL
LVKHADRNLTIVQQAQVTVANGNKRKSNLDEGHALRIPSGFISYILNRHQN
LRVAKIAMPNTPGQEDFFPASRDSSSYLQFSRNTLEAANFAEFNEIRHJLEEN
AGGQBERGQRRSTRSDNKGVIKVSKEHVBLTHAKSVSKGSEEDITNPIL
RDGPDLSSNFGRLFEVKPKDNQLOLDMLTCEIKEGALMLPHNSKAMVTVV
NKGITNLELVAVRKEOQGRREGEWEEREEDEEGSNRVRVYTABLKEGDVIMP
AAHPVNASSELHLLFGGINAENHRIFLAGDKDNVIDQIEKQAKDLAFPGSGEQVE
KLINQKRESHFVSA"

ORIGIN

Query Match 68.7%; Score 1395.4; DB 8; Length 2332;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 46; Indels 253; Gaps 8;
QY 2 ATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAGG 61
Db ATAATAATACATATATTCATCAATCATCTATATATAAGTAGTTGCAGGAGCAATGAGAGGAGG 354
QY 62 GTTTCCTCACTGATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCAT 121
Db GTTTCCTCACTGATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCAG 414
QY 122 GCCAAGTCATCACTTACAGAGAAAGAAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCAG 181
Db GCCAAGTCACCC-----TTACCGGNAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCAG 468
QY 182 AGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGTCTCGTGTGCAACCAAG 241
Db AGTTGTCAACAGGAACCGGACGACTTGAAGCAAAAGGATGCGAGTCTCGTGTGCAACCAAG 528
QY 242 CTCGAGTATGATCTCTGTTGTCTATGATCCTCGAGGACACACTGCGCACCAACCAACCA 301
Db CTCGAGTATGATCTCTGTTGTGTCTATG-----ACACTGGCGCCACCAACCA 576
QY 302 CGTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCCGGAGACTAGCATGATGACCGC 361
Db CGTCACCTCCAGGGAGCGGACACGTCGCGCCCAACCCCGGAGACTAGCATGATGACCGC 636
QY 362 CGTCAACCCCGAAGAGAGGAGCGCGATGGGGAACAGCTGGACCGAGGAGCGTGAA 421
Db CGTCAACCCCGAAGAGAGGAGCGCGATGGGGAACAGCTGAACCCGAGGAGCGTGAA 696
QY 422 AGAGAAAGAGCTGGAGACACCAAGAGAGATTGGAGCGGACCAAGTTCATCAGCAGCA 481
Db AGAGAAAGAGCTGGAGACACCAAGAGAGATTGGAGCGGACCAAGTTCATCAGCAGCA 756
QY 482 CGGAAATTAAGCGCCGAAGGAAGAGAGAAACAAGAGTGGGGAACAACCAAGGTAGCCAT 541
Db CGGAAATTAAGCGCCGAAGGAAGAGAGAAACAAGAGTGGGGAACAACCAAGGTAGCCAG 816
QY 542 GTGAGGGAAGAAACATCTCGGAAACAACCTTTCTATCTTCCTGCAAGCGGTTTAGCACC 601
Db GTGAGGGAAGAGACATCACGGNAACAACCTTTCTACTTCCGCTCAAGCGGTTTAGCACC 876
QY 602 CGCTACGGGAACCAAAACGGTAGGATCGGGTCTCTGCAAGAGTTTGACCAAGGTCAAGG 661
Db CGCTACGGGAACCAAAACGGTAGGATCGGGTCTCTGCAAGAGTTTGACCAAGGTCAAGG 936
QY 662 CAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGAGATCGAGGCGCAAACTTAACTCTT 721
Db CAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGAGATCGAGGCGCAAACTTAACTCTT 996
QY 722 GTTCTTCCCAAGCAGCTGTGATGAATAACATCTCTGTTTATTCACAGCA----- 768
Db GTTCTTCCCAAGCAGCTGTGATGAATAACATCTCTGTTTATTCACAGCAAGGTATCAATCT 1056
QY 769 -----AG 770
Db AATCTATCTTACTATATATATATTTTGTGTTGATACATATGATTCATGATTGCGAG 1116

Qy	771	GGCAAGCCACCGTGACCGTACCAATATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAGG	830
Db	1117	GACAAGCCACCGTGACCGTACCAATATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAGG	1176
Qy	831	GCCATGCACCTCAGAAATCCCATCCGGTTTCATTTCTCCTACATCTTGAAACGCGCATGACAAAC	890
Db	1177	GCCATGCACCTCAGAAATCCCATCCGGTTTCATTTCTCCTACATCTTGAAATCGCATGACAAAC	1236
Qy	891	AGAACTCTCAGAGTAGCTAAAATCTCCATGCCCGTTTAAACACACCCCGCGCAGTTTG-	944
Db	1237	AGAACTCTCAGAGTAGCTAAAATCTCCATGCCCGTTTAAACACACCCCGCGCAGTTTGAGGTAG	1296
Qy	945	-----	944
Db	1297	GTACCTCTTTCTCTCACATATATATTAATCTCAATATATCATCTTTACATGTTGTGGT	1356
Qy	945	-----AGGATTTCTCCCGCGAGCAGCGAGACCAATCATCTACTTTCAGGGC	994
Db	1357	GTTGCTTACAGGATTTCTCCCGCGAGCAGCGAGACCAATCATCTACTTTCAGGGGA	1416
Qy	995	TTACAGCAGGAATACGTTTGGAGCGCGCTTCAAT	1027
Db	1417	TTACAGCAGGAATACTTTTGGAGCGCGCTTCAATGTAGAACAAATGTGTCAATAATTATGGAA	1476
Qy	1028	-----	1041
Db	1477	TTAAAGAAACGATCATGTATTAATCTTATAATATATATATACATAGCGGAATTCATGA	1536
Qy	1042	GATACGAGGGTGTCTTTAGNAGAAATGACGAGGTGACGACGAGAGAGAGGGCAGAG	1101
Db	1537	GATACGAGGGTGTCTTTAGNAGAAATGACGAGGAGAGCAGAGAGAGAGAGGGCAGAG	1596
Qy	1102	GCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTGATAGTCAAAGTGTCAAAGGA	1161
Db	1597	GCGACGGAGTACTCGGAGTAGTGAT--AATGAAGAGTAGTAGTCAAAGTGTCAAAGGA	1653
Qy	1162	GCACGTTGAAGAACTTACTAAGCACGCTAAATCCGTCTCAAAGAAAGGCTCCCGAAGAAGA	1221
Db	1654	GCACGTTCAAGAACTTACTAAGCACGCTAAATCCGTCTCAAAGAAAGGCTCC--GAAGA	1710
Qy	1222	GGGAGATATCAACCAATCACTTTGAGAGAGGCGAGCCGATCTTTCTAACCACTT	1281
Db	1711	GGAAGATATCAACCAATCACTTTGAGAGATGGCGAGCCGATCTTTCTAACCACTT	1770
Qy	1282	TGGGAAGTATTTGAGGTGAAGCCAGACAAGAAGACCCCGCTTCAGGACCTGGACAT	1341
Db	1771	TGGGAGGTATTTGAGGTGAAGCCAGACAAGAAGACCCCGCTTCAGGACCTGGACAT	1830
Qy	1342	GATGCTCACCTGTGTAGAGATCAAAGAGAGCTTTGATGCTCCACATCTCAACTCAAA	1401
Db	1831	GATGCTCACCTGTGTAGAGATCAAAGAGAGCTTTGATGCTCCACATCTCAACTCAAA	1890
Qy	1402	GGCCATGGTTATCGTCTGTCTCAACAAAGAACTGGAAACTCTGGAACCTCGTGCTGTAAAG	1461
Db	1891	GGCCATGGTTCATCGTCTGTCTCAACAAAGAACTGGAAACTCTGGAACCTCGTGCTGTAAAG	1950
Qy	1462	AAAAGACCAACACAGAGGGGACGGCGGAA-----GAAGAGGAGACCAAGACGA	1512
Db	1951	AAAAGACCAACACAGAGGGGACGGCGGAAACAGAGTGGGAAGAGAGGAGGAAGATGA	2010
Qy	1513	AGAAGAGGAGGAGTAAACAGAGAGGTGCGTAGGTACAACGAGAGGTTGAAGGAAGCGCA	1572
Db	2011	AGAAGAGGAGGAGTAAACAGAGAGGTGCGTAGGTACAACGAGAGGTTGAAGGAAGCGCA	2070
Qy	1573	TGTGTTTATATGCGCAGAGCTCATTCAGTAGCCATCAACGCTTCTCCGAACTCCATCT	1632
Db	2071	TGTGTTTATATGCGCAGAGCTCATTCAGTAGCCATCAACGCTTCTCCGAACTCCATCT	2130
Qy	1633	GCTTGGCTTCGGTATCAACGCTGAACCAACCAACAGAAATCTTCTTTCGAGGTGATGAAGA	1692
Db	2131	GCTTGGCTTCGGTATCAACGCTGAACCAACCAACAGAAATCTTCTTTCGAGGTGATGAAGA	2190
Qy	1693	CAATGTGATAGACCAGATAGAGAGCAAGCAAGGAAGGATTTTAGCATTCCTCGGTCGGGTGA	1752

Db	2191	CAATGTGATAGACAGATAGAGAACGAAGGATTAGCATTCCTCGTTCGGGTGA	2250
Qy	1753	ACAAAGTTTGAGAGGCTCATCAAAACACGAGAGGAATCTCCTTTGTGAGTGCTCGTCTCTCA	1812
Db	2251	ACAAAGTTTGAGAGGCTCATCAAAACACGAGAGGAGTCTCCTTTGTGAGTGCTCGTCTCTCA	2310
Qy	1813	ATCTCAATCTCAATC	1827
Db	2311	ATCTCAATCTCGTC	2325
RESULT 14			
LOCUS	AY581851	1418 bp	linear
DEFINITION	Arachis hypogaea conarachin mRNA, partial cds.		PLN 01-MAY-2004
ACCESSION	AY581851		
VERSION	AY581851.1	GI:46560475	
KEYWORDS			
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
REFERENCE	1 (bases 1 to 1418)		
AUTHORS	Wang,L., Liao,B., Li,H., Yan,Y., Lin,X. and Huang,S.		
TITLE	cDNA cloning of conarachin in peanut seed		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1418)		
AUTHORS	Wang,L., Liao,B., Li,H., Yan,Y., Lin,X. and Huang,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAR-2004) Department of Biology and Biotechnology, School of Life Science, Sun Yat-sen University, Guangzhou, Guangdong 510275, China		
FEATURES	Location/Qualifiers		
source	1..1418		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/cultivar="SanYou523"		
	/db_xref="taxon:3818"		
	<1..1289		
	/note="similar to peanut allergen 1 (p41b)"		
	/codon_start=3		
	/product="conarachin"		
	/protein_id="AAT00596.1"		
	/db_xref="GI:46560476"		
	/translation="GPDQRSPQNLQNHRIVOIEAKPNTLVLPKHADADNLIVIOGEOATVTVANGNRKSFNLDGHALRIPSGFTSYILNRHDNLRVAKISMPVNTPGOFEDFPASRDSSVYQFSRNTLEAANFNEIRRVLEENAGEDEQRQSRWSFTRSENNEGVIVKSVHEELTKHAKSVKSGSEBDITNINLRGEGFDLSNNTFGKLFVVKPDKNPQODLMDMLTCVEIKEGALMPLHFENSKAMVIVVKNKGTDNLULVAVRKQQRGRREEDDEEGNSRVRYRTALPGKGDVFIMPAAHFVAINASELHLGFGINAENNHRIFLAGDKNVLDQIEKQAKDLAFPGSGEQVEKLIKQKESHFVSARPSQSQSPSSPEKSPESPEKDOEEENQGGKPLLSILKAFN"		
CDS			
ORIGIN			
Query Match	67.7%	Score 1375.4;	DB 8; Length 1418;
Best Local Similarity	99.7%	Pred. No. 0;	
Matches 1389;	Conservative	0; Mismatches	1; Indels 3; Gaps 1;
Qy	640	GAGGTTTGACCAAGGTCAGGAGGTTTCAGATCTCCAGATCCACCGTATTCGCGAGAT	699
Db	2	GGGGTTTGACCAAGGTCAGGAGGTTTCAGATCTCCAGATCCACCGTATTCGCGAGAT	61
Qy	700	CGAGGCCAAACCTAAACACTCTTGTTCTTCCCAAGCAGCTGATGCTGATTAACATCTCTTGT	759
Db	62	CGAGGCCAAACCTAAACACTCTTGTTCTTCCCAAGCAGCTGATGCTGATTAACATCTCTTGT	121
Qy	760	TATCCAGCAGGCGAACGCCACCGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAA	819
Db	122	TATCCAGCAGGCGAACGCCACCGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAA	181

QY	820	TCTTGACGAGGCCATGCACTCAGAAATCCCATCCGGTTTCATTTCCTCATCTTTGAACCG	879
Db	182	TCTTGACGAGGCCATGCACTCAGAAATCCCATCCGGTTTCATTTCCTCATCTTTGAACCG	241
QY	880	CCATGACACACGAACTCAGAGTAGCTAAAATCTCCATGCCGGTTAAACACACCCGGCCA	939
Db	242	CCATGACACACGAACTCAGAGTAGCTAAAATCTCCATGCCGGTTAAACACACCCGGCCA	301
QY	940	GTTTGGAGGATTCTTCCCGCGCAGACGCGAGACCAATCATCTTCTTCACGGCTTCAG	999
Db	302	GTTTGGAGGATTCTTCCCGCGCAGACGCGAGACCAATCATCTTCTTCACGGCTTCAG	361
QY	1000	CAGGAATACGTTGGAGGCCCTTCAATCCGGAATTCATAGATACGAGGGTCTGTT	1059
Db	362	CAGGAATACGTTGGAGGCCCTTCAATCCGGAATTCATAGATACGAGGGTCTGTT	421
QY	1060	AGAAGAGATCAGAGGTTGACCAAGAGAGAGAGGGCGAGCGATGGAGTACTCGGAG	1119
Db	422	AGAAGAGATCAGAGGTTGACCAAGAGAGAGAGGGCGAGCGATGGAGTACTCGGAG	481
QY	1120	TAGTGAGAACAAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGTTTGAAGAACTTAC	1179
Db	482	TAGTGAGAACAAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGTTTGAAGAACTTAC	541
QY	1180	TAAGCACGCTAAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACC	1239
Db	542	TAAGCACGCTAAATCCGTCTCAAAGAAAGGCTCC--GAAGAGGAGATATCACCAACC	598
QY	1240	AATCAACTTGAGAGAGGCGCGATCTTTCTAACAACTTTGGGAAGTTATTGGAGT	1299
Db	599	AATCAACTTGAGAGAGGCGCGATCTTTCTAACAACTTTGGGAAGTTATTGGAGT	658
QY	1300	GAAGCAGACAAAGAAACCCCGCTTCAGACCTGGACATGATGCTCACCTGTGTAGA	1359
Db	659	GAAGCAGACAAAGAAACCCCGCTTCAGACCTGGACATGATGCTCACCTGTGTAGA	718
QY	1360	GATCAAGAGAGCTTTGATGCTCCACATTCAACTCAAAGGCCATGTTTATCGTGT	1419
Db	719	GATCAAGAGAGAGCTTTGATGCTCCACATTCAACTCAAAGGCCATGTTTATCGTGT	778
QY	1420	CGTCAACAAAGAACTGGNAACTTGAACTCGTGGCTGTAGAAAAGAGCAACACAGAG	1479
Db	779	CGTCAACAAAGAACTGGNAACTTGAACTCGTGGCTGTAGAAAAGAGCAACACAGAG	838
QY	1480	GGGACGGCGGGAAGAGAGGACGACGAAGCAAGAGAGGAGGAAGTAAACAGAGAGT	1539
Db	839	GGGACGGCGGGAAGAGAGGACGACGAAGCAAGAGAGGAGGAAGTAAACAGAGAGT	898
QY	1540	GGGTAGGTACACAGCGAGGTTGAAGGAAGGGCGATGTGTTTCATGATGCCAGCAGCTCATCC	1599
Db	899	GGGTAGGTACACAGCGAGGTTGAAGGAAGGGCGATGTGTTTCATGATGCCAGCAGCTCATCC	958
QY	1600	AGTAGCCATCAACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAA	1659
Db	959	AGTAGCCATCAACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAA	1018
QY	1660	CAACACACAGAAATCTTCTTGACAGGTGATAAGGACAAATGTGTAGACCAAGATAGAGAACA	1719
Db	1019	CAACACACAGAAATCTTCTTGACAGGTGATAAGGACAAATGTGTAGACCAAGATAGAGAACA	1078
QY	1720	AGCGAAGGATTAGCATTCCTCGGTTCGGGTGAAACAAAGTTGAGAGCTCATCAAAAAACA	1779
Db	1079	AGCGAAGGATTAGCATTCCTCGGTTCGGGTGAAACAAAGTTGAGAGCTCATCAAAAAACA	1138
QY	1780	GAAGGAATCTCACTTTGTGAGTGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCC	1839
Db	1139	GAAGGAATCTCACTTTGTGAGTGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCC	1198
QY	1840	TGAGAAAGAGTCTCTTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGGAAGGGTCC	1899
Db	1199	TGAGAAAGAGTCTCTTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGGAAGGGTCC	1258
QY	1900	ACTCCTTTCAAATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTGTATGATCGATA	1959

Db	1259	ACTCCTTTCAAATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTGTATGATCGATA	1318
QY	1960	ATAAGATCAGCGTTTGTACTCTACTATCCAAAACTTATCAATAAATAAACAAGCTTTGT	2019
Db	1319	ATAAGATCAGCGTTTGTACTCTACTATCCAAAACTTATCAATAAATAAACAAGCTTTGT	1378
QY	2020	CGTTGTTTCTCC	2032
Db	1379	CGTTGTTTCTCC	1391
RESULT 15	AY581850	1035 bp mRNA linear	PLN 01-MAY-2004
LOCUS	Arachis hypogaea conarachin mRNA, partial cds.		
DEFINITION	Arachis hypogaea		
ACCESSION	AY581850		
VERSION	AY581850.1	GI:46560473	
KEYWORDS			
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
AUTHORS	Wang, L., Liao, B., Li, H., Yan, Y., Lin, X. and Huang, S.		
TITLE	CDNA cloning of conarachin in peanut seed		
JOURNAL	Unpublished		
AUTHORS	Wang, L., Liao, B., Li, H., Yan, Y., Lin, X. and Huang, S.		
TITLE	Direct Submision		
JOURNAL	Submitted (01-MAR-2004) Department of Biology and Biotechnology, School of Life Science, Sun Yat-sen University, Guangzhou, Guangdong 510275, China		
FEATURES	Location/Qualifiers		
source	1..1035		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/cultivar="SanYou523"		
	/db_xref="taxon:3818"		
	<1..901		
	/note="similar to peanut allergen 1 (p17)"		
	/codon_start=2		
	/product="conarachin"		
	/protein_id="AA00595.1"		
	/db_xref="GI:46560474"		
	/translation="LEAFAFNEIRRVLEENAGGEQERGRWRSTRSENNEGV IVKYSKHVEBELTKHAKSVSKGSEEGDITNPINLRGEPDLNNFKLFEVQPKK NPOQLDMLTCEIKEGALMLPHFNKAMVI VVNVKGTGNLSELVAVRKEQQRGR EEEDEQDEBGSNRVRYTARLKEGDFVIMPAHPVAI NASSELHLGFGINAENN HRIPLAGDKNVIDQIEKQAKDLAPGSGEQVEKLKNQRESHFVSARPOQSPPSE KEDDEENQGGKGFLLSKAFN"		
CDS			
ORIGIN			
Query Match	47.0%;	Score 954.4;	DB 8;
Best Local Similarity	96.9%;	Pred. No. 2.3e-258;	
Matches	992;	Conservative 0;	Mismatches 11;
		Indels 21;	Gaps 1;
QY	1009	GTTGAGGCCCTTCAATCGGAATTCATAGATACGAGGGTGTGTTAGAGAGAA	1069
Db	1	GTTGAGGCCCTTCAATCGGAATTCATAGATACGAGGGTGTGTTAGAGAGAA	60
QY	1069	TGCAGGAGTGCAGCAGAGGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAAA	1128
Db	61	TGCAGGAGTGCAGCAGAGGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAAA	120
QY	1129	CAATCAAGAGTGTAGTCAAAAGTGTCAAAGCAGCAGCTTCAAGAACTTACTTAAGCAGC	1188
Db	121	CAATCAAGAGTGTAGTCAAAAGTGTCAAAGCAGCAGCTTCAAGAACTTACTTAAGCAGC	180
QY	1189	TAAATCGTCTCAAAGAAAGGCTCCGAGAAAGAGGAGATATCACCAACCAATCACTT	1248


```
Db      181  TAAATCCGTCCTCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCAACTT 240
Qy      1249  GAGAGAAGCGAGCCCGATCTTTCTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGA 1308
Db      241  GAGAGAAGCGAGCCCGATCTTTCTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGA 300
Qy      1309  CAAGAAGAACCCCGACGCTTCAGAGCCTGGACATGATGCTCACCTGTGTAGAGATCAAGA 1368
Db      301  CAAGAAGAACCCCGACGCTTCAGAGCCTGGACATGATGCTCACCTGTGTAGAGATCAAGA 360
Qy      1369  AGGAGCTTTGATGCTCCACACCTTCAACTCAAAGGCCATGGTTATCGTCGTCTCAACAA 1428
Db      361  AGGAGCTTTGATGCTCCACACCTTCAACTCAAAGGCCATGGTTATCGTCGTCTCAACAA 420
Qy      1429  AGAACTGGAAACCTTGAACTGTGTGCTGTAGAAAGCAACAACAGAGGGGACGGCG 1488
Db      421  AGGAACTGGAAACCTTGAACTGTGTGCTGTAGAAAGCAACAACAGAGGGGACGGCG 480
Qy      1489  GGAAGAAGAGGAGGACGAAAGAGAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTA 1548
Db      481  GGAAGAAGAGGAGGACGAAAGAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTA 540
Qy      1549  CACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCAT 1608
Db      541  CACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCAT 600
Qy      1609  CAAAGCTTCTCCGAACCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668
Db      601  CAAAGCTTCTCCGAACCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      1669  AATCTTCTTGCAGGTGATAGGACAACTGTGTAGACCAAGATAGAGAGGAGGAGGAGGAG 1728
Db      661  AATCTTCTTGCAGGTGATAGGACAACTGTGTAGACCAAGATAGAGAGGAGGAGGAGGAG 720
Qy      1729  TTTAGCATTTCCCTGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGGAATC 1788
Db      721  TTTAGCATTTCCCTGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGGAGTC 780
Qy      1789  TCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCAATCTCCTGCTGCTGCTGCTGCTGCT 1848
Db      781  TCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 819
Qy      1849  GTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1908
Db      820  GTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 879
Qy      1909  AATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAGATCA 1968
Db      880  AATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAGATCA 939
Qy      1969  CGCTTTGTACTTACTATCCAAAACTTATCAATAAATAAATAAATAAATAAATAAATAAATA 2028
Db      940  CGCTTTGTAACTACTATCCAAAACTTATCAATAAATAAATAAATAAATAAATAAATAAATA 999
Qy      2029  CTCC 2032
Db      1000  CTCC 1003
```

Search completed: August 24, 2005, 00:57:06
Job time : 6030.11 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 4765.9 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataacatatattcatc.....cgttgtggtgtttcttc 2032

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	719	35.4	721	6	CD038628
2	713	35.1	714	6	CD038277
3	711.4	35.0	723	6	CD038527
4	681	33.5	717	6	CD038837
5	657.2	32.3	724	6	CD038648
6	629.2	31.0	719	6	CD038172
7	626	30.8	676	6	CD038790
8	602	29.6	685	6	CD038555
9	576	28.3	588	6	CD038119
10	527.6	26.0	684	6	CD038540
11	473	23.3	474	6	CD038694
12	471	23.2	582	6	CD038620
13	458.8	22.6	509	6	CD038253
14	436	21.5	484	6	CD038101
15	428.6	21.1	443	7	CO897502
16	425	20.9	425	6	CD038284
17	419.6	20.6	536	6	CD038075
18	392	19.3	411	7	CO897503
19	358.6	17.6	406	6	CD038765
20	307.6	15.1	823	6	CA858229
21	306.2	15.1	851	6	CA858769
22	296.8	14.6	894	6	CA858664
23	296.6	14.6	806	4	BI310422
24	292.2	14.4	864	5	BQ123471

25	290.8	14.3	780	4	BI311149
26	290.6	14.3	823	5	BQ122514
27	290	14.3	862	5	BQ123135
28	285.8	14.1	854	5	BQ123411
29	285.4	14.0	830	6	CA858151
30	283	13.9	805	4	BI312339
31	281.4	13.8	760	6	CA858137
32	281.4	13.8	769	4	BI311022
33	280.2	13.8	847	6	CA858293
34	278.4	13.7	781	4	BI311613
35	277.2	13.5	765	5	BQ122510
36	273.2	13.4	804	6	CA858167
37	272.4	13.4	843	4	BI311012
38	271.8	13.4	805	6	CA858433
39	271.4	13.4	734	6	CA858790
40	271.4	13.4	839	4	BI311235
41	271.2	13.3	861	6	CA858864
42	271	13.3	860	5	BQ122453
43	270.4	13.3	701	5	BQ124840
44	270.4	13.3	727	4	BI309850
45	270.4	13.3	837	4	BI310945

ALIGNMENTS

RESULT 1
CD038628
LOCUS CD038628 721 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI010_A05 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI010_A05 5', mRNA sequence.
ACCESSION CD038628
VERSION CD038628.1 GI:30420466
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 721)
AUTHORS Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G. and Lynch, R.E.
TITLE Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL Unpublished (2003)
COMMENT Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES
source

Location/Qualifiers
1..721
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI010_A05"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCVILXAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapur (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN		Query Match	35.4%;	Score 719;	DB 6;	Length 721;
		Best Local Similarity	99.7%;	Pred. No. 4.3e-191;		
		Matches 719;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1231	CACCAACCCCAATCAACTTGAGAGAGCGCGAGCCGATCTTTCTAACAACCTTTGGGAAGTT	1290			
Db	1	CACCAACCCCAATCAACTTGAGAGAGCGCGAGCCGATCTTTCTAACAACCTTTGGGAAGTT	60			
Qy	1291	ATTTGAGGTGAAGCCAGACAGAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCAC	1350			
Db	61	ATTTGAGGTGAAGCCAGACAGAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCAC	120			
Qy	1351	CTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACCTTCAACTCAAAAGGCCATGTT	1410			
Db	121	CTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACCTTCAACTCAAAAGGCCATGTT	180			
Qy	1411	TATCGTCGTCTCAACAAGAGAACTGGAAACCTTGAACCTTGGCTGTGAAGAAAGAGCA	1470			
Db	181	TATCGTCGTCTCAACAAGAGAACTGGAAACCTTGAACCTTGGCTGTGAAGAAAGAGCA	240			
Qy	1471	ACAACAGAGGGGACGGCGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTAA	1530			
Db	241	ACAACAGAGGGGACGGCGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTAA	300			
Qy	1531	CAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAGGCGATGTGTTTCATATGCGCAGC	1590			
Db	301	CAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAGGCGATGTGTTTCATATGCGCAGC	360			
Qy	1591	AGTCTATCCAGTAGCCATCAACCGTTTCTCCGAACTCCCATCTGCTTGGCTTGGTATCAA	1650			
Db	361	AGTCTATCCAGTAGCCATCAACCGTTTCTCCGAACTCCCATCTGCTTGGCTTGGTATCAA	420			
Qy	1651	CGCTGAACCAACACAGAGATCTTCTTCGAGGTGATAGGACAATGTGATAGACCAGAT	1710			
Db	421	CGCTGAACCAACACAGAGATCTTCTTCGAGGTGATAGGACAATGTGATAGACCAGAT	480			
Qy	1711	AGAGAAGCAAGCGAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAAGCTCAT	1770			
Db	481	AGAGAAGCAAGCGAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAAGCTCAT	540			
Qy	1771	CAAAAACCGAAGGAATCTCATCTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCC	1830			
Db	541	CAAAAACCGAAGGAATCTCATCTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCC	600			
Qy	1831	GTGCTCTCTGAGAAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGG	1890			
Db	601	GTGCTCTCTGAGAAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGG	660			
Qy	1891	GAAGGGTCCACTCTCTTTCAATTTTGAAGGCTTTTAACTGAGAATGGAGGCACTTCTTAT	1950			
Db	661	GAAGGGTCCACTCTCTTTCAATTTTGAAGGCTTTTAACTGAGAATGGAGGCACTTCTTAT	720			
Qy	1951	G 1951				
Db	721	G 721				
RESULT 2		CD038277	714 bp	mRNA	linear	EST 07-MAY-2003
LOCUS						

UTPPI005 A03 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI005_A03 5', mRNA sequence.
CD038277
CD038277.1 GI:30420115
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 714)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..714
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI005_A03"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV11XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIZOL-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match		35.1%;	Score 713;	DB 6;	Length 714;
Best Local Similarity		99.9%;	Pred. No. 2.1e-189;		
Matches 713;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	146	AAAAACAGAGAACCCCTCGCGCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGAC	205		
Db	1	AAAAACAGAGAACCCCTCGCGCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGAC	60		
Qy	206	TTGAAGCAAAAGGCATGGAGTCTCGTCGACCAAGATCGAGTATGATCTCTGTGTGTC	265		
Db	61	TTGAAGCAAAAGGCATGGAGTCTCGTCGTCGACCAAGATCGAGTATGATCTCTGTGTGTC	120		
Qy	266	TATGATCTCTGAGGACACACTGGCCACCAACCAAGCTTCCCTCCAGGAGGCGGACA	325		
Db	121	TATGATCTCTGAGGACACACTGGCCACCAACCAAGCTTCCCTCCAGGAGGCGGACA	180		

```
Qy 326 CGTGGCCGCAACCCGAGACTACGATGATGACCGCGTCAACCCCGAAGAGGAGGA 385
Db 181 CGTGGCCGCAACCCGAGACTACGATGATGACCGCGTCAACCCCGAAGAGGAGGA 240
Qy 386 GGCCGATGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAGAAAGACTGGAGACACCA 445
Db 241 GGCCGATGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAGAAAGACTGGAGACACCA 300
Qy 446 AGAGAAGATTGGAGGCGACCAAGTTCATCAGCAGCCACGGAAATTAAGCCCGAAGGA 505
Db 301 AGAGAAGATTGGAGGCGACCAAGTTCATCAGCAGCCACGGAAATTAAGCCCGAAGGA 360
Qy 506 GAAGGAGACAAGAGTGGGAGCAACACAGTAGTACCATGTGAGGAGAAAGATCTCGCAAC 565
Db 361 GAAGGAGACAAGAGTGGGAGCAACACAGTAGTACCATGTGAGGAGAAAGATCTCGCAAC 420
Qy 566 AACCTTTTCTACTTCCCGTCAAGCGCGTTTACGACCCGCTACGGGAACCAAAACGGTAGG 625
Db 421 AACCTTTTCTACTTCCCGTCAAGCGCGTTTACGACCCGCTACGGGAACCAAAACGGTAGG 480
Qy 626 ATCCGGTCTCTGACAGGTTTGACCAAGGTCAAGGAGGTTTCAAGATCTCCGAATCAC 685
Db 481 ATCCGGTCTCTGACAGGTTTGACCAAGGTCAAGGAGGTTTCAAGATCTCCGAATCAC 540
Qy 686 CGTATTGTGAGATCGAGGCGCAACCTAACACTCTTCTTCTCCAGACGCTGATGCT 745
Db 541 CGTATTGTGAGATCGAGGCGCAACCTAACACTCTTCTTCTCCAGACGCTGATGCT 600
Qy 746 GATAACATCTTGTATTCAGCAGGCGCAAGCCACCGTAGCCGTAAGAAATGGCAATAAC 805
Db 601 GATAACATCTTGTATTCAGCAGGCGCAAGCCACCGTAGCCGTAAGAAATGGCAATAAC 660
Qy 806 AGAAGAGCTTTAATCTTGACGAGGCGCATGCACTCAGATCCCATCCGTTTC 859
Db 661 AGAAGAGCTTTAATCTTGACGAGGCGCATGCACTCAGATCCCATCCGTTTC 714

RESULT 3
CD038527 723 bp mRNA linear EST 07-MAY-2003
LOCUS UTPI008_E12 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION Arachis hypogaea cDNA clone UTPI008_E12 5', mRNA sequence.
ACCESSION CD038527
VERSION CD038527.1 GI:30420365
SOURCE EST.
ORGANISM Arachis hypogaea (peanut)
            Arachis hypogaea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
            Aeschynomeneae; Arachis.
            1 (bases 1 to 723)
            Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
            and Lynch, R. E.
            Generation and Analyses of ESTs for Arachis hypogaea
            Unpublished (2003)
            Contact: Baozhu Guo
            Molecular Genetics
            USDA/ARS, Crop Protection and Management Research Unit
            2747 Davis Rd., Tifton, GA 31794, USA
            Tel: 229-387-2334
            Fax: 229-387-2321
            Email: bguo@tifton.usda.gov
            Seq primer: T3.
            Location/Qualifiers
            1..723
            /organism="Arachis hypogaea"
            /mol_type="mRNA"
            /cultivar="A13"
            /db_xref="taxon:3818"
            /clone="UTPI008_E12"
            /tissue_type="Immature pods"
            /dev_stage="R6"

FEATURES
            source
```

```
/lab_host="XLI-blue"
/clone_1lb="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV11XA4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

ORIGIN

Query Match	35.0%	Score 711.4	DB 6	Length 723
Best Local Similarity	99.4%	Pred. No. 6e-189		
Matches 712	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy 1210	CTCCGAAGAGAGGAGATATACCAACCCAACTCAACTTGAGAGAGGCGAGCCGATCT	1269		
Db 1	CTCCGAAGAGAGGAGATATACCAACCCAACTCAACTTGAGAGAGGCGAGCCGATCT	60		
Qy 1270	TTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCCCGATTCA	1329		
Db 61	TTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCCCGATTCA	120		
Qy 1330	GGACCTGGACATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCACA	1389		
Db 121	GGACCTGGACATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCACA	180		
Qy 1390	CTTCAACTCAAGGCCATGGTTATCGTCGTCAACAAAGGAACCTGAAACCTTGAAC	1449		
Db 181	CTTCAACTCAAGGCCATGGTTATCGTCGTCAACAAAGGAACCTGAAACCTTGAAC	240		
Qy 1450	CGTGGCTGTAAGAAAAGAGCAACAACAGAGGGGACCGCGGAAGAGAGGAGCAAGA	1509		
Db 241	CGTGGCTGTAAGAAAAGAGCAACAACAGAGGGGACCGCGGAAGAGAGGAGCAAGA	300		
Qy 1510	CGAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAGGAGG	1569		
Db 301	CGAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAGGAGG	360		
Qy 1570	CGATGTTTCATCATGCCAGCCTCATCCAGTAGGCATCAACGCTTCCTCCGAACCTCA	1629		
Db 361	CGATGTTTCATCATGCCAGCCTCATCCAGTAGGCATCAACGCTTCCTCCGAACCTCA	420		
Qy 1630	TTCTGTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATA	1689		
Db 421	TTCTGTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATA	480		
Qy 1690	GGCAATGTGATAGACAGATAGAGAGCAGAGGATTTAGCATTCCTCGGTCGGG	1749		
Db 481	GGCAATGTGATAGACAGATAGAGAGCAGAGGATTTAGCATTCCTCGGTCGGG	540		
Qy 1750	TGAACAAGTTGAGAAGCTCATCAAAAACAGAGAGGAATCTCACTTTGTAGTGTCTGTC	1809		
Db 541	TGAACAAGTTGAGAAGCTCATCAAAAACAGAGAGGAATCTCACTTTGTAGTGTCTGTC	600		
Qy 1810	TCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTGAGAAAGAGGATCA	1869		
Db 601	TCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCN	660		
Qy 1870	AGAGGAGGAAAACCAAGGAGGAGGGTCCACTCTCTTCAATTTTGNAGGCTTTTA	1925		

```
Db 661 AGAGGAGAAACGAGGGGAGGGTCCACTCTTCNTATTTTGAAGGCTTTTA 716
|||||
RESULT 4
CD038837 717 bp mRNA linear EST 07-MAY-2003
LOCUS UTTP1012 H04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION Arachis hypogaea cDNA clone UTTP1012_H04 5', mRNA sequence.
ACCESSION CD038837
VERSION 1 GI:30420675
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 717)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
FEATURES
Location/Qualifiers
1..717
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTTP1012 H04"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="X11-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar Al3
(NCV11XR4). Al3 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."
ORIGIN
Query Match 33.5%; Score 681; DB 6; Length 717;
Best Local Similarity 99.6%; Pred. No. 2.3e-180;
Matches 714; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 967 CCGAGACCAATCATCTTACTTGACGGCTTCACAGGAATACGTTGGAGCCGCTTCAA 1026
|||||
Db 1 CCGAGACCAATCATCTTACTTGACGGCTTCACAGGAATACGTTGGAGCCGCTTCAA 60
|||||
QY 1027 TCGGGAATTCATGAGATACGGAGGGTGTGTTTGAAGAAGATGCGAGGAGGTGAGCAAGA 1086
|||||
```

```
Db 61 TCGGGAATTCATGAGATACGGAGGGTGTGTTTGAAGAAGATGCGAGGAGGTGAGCAAGA 120
|||||
QY 1087 GGAGAGAGGCGAGAGCGGATGCGAGTACTCGGAGTACTGAGAACATGAAGGAGTGTAGT 1146
|||||
Db 121 GGAGAGAGGCGAGAGCGGATGCGAGTACTCGGAGTACTGAGAACATGAAGGAGTGTAGT 180
|||||
QY 1147 CAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTCAAAGAA 1206
|||||
Db 181 CAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTCAAAGAA 240
|||||
QY 1207 AGGCTCCGGAAGAAGAGGGAGATATCAACCAATCAACTTTGAGAGAAGGCGAGCCGA 1266
|||||
Db 241 AGGCTCCGGAAGAAGAGGGAGATATCAACCAATCAACTTTGAGAGAAGGCGAGCCGA 300
|||||
QY 1267 TCTTTTCTAACTTTGGGAAGTATTTGAGGTGAAGCCAGCAAGCAAGCAAGCAAGCCAGCT 1326
|||||
Db 301 TCTTTTCTAACTTTGGGAAGTATTTGAGGTGAAGCCAGCAAGCAAGCAAGCAAGCCAGCT 360
|||||
QY 1327 TCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCC 1386
|||||
Db 361 TCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCC 420
|||||
QY 1387 ACACCTCAACTCAAAGGCCATGTTATCGTCTGTCAACAAAGGAACCTGGAACCTTGA 1446
|||||
Db 421 ACACCTCAACTCAAAGGCCATGTTATCGTCTGTCAACAAAGGAACCTGGAACCTTGA 480
|||||
QY 1447 ACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGCGCGGGAAGAGAGGAGGACGA 1506
|||||
Db 481 ACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGCGCGGGAAGAGAGGAGGACGA 540
|||||
QY 1507 AGACGAAGAAG-AGGAGGGAAGTAACAGA-GAGGTGCGTAGGTACACACGCGAGTTGAAG 1564
|||||
Db 541 AGACGAAGAAGAGGAGGGAAGTAACAGAGGAGGTGCGTAGGTACACACGCGAGTTGAAG 600
|||||
QY 1565 GAAGCGGATGTTTATCATATGCCAGCAGCTATCAGTAGCCATCAACGCTTCTCTCCGA 1624
|||||
Db 601 GAAGCGGATGTTTATCATATGCCAGCAGCTATCAGTAGCCATCAACGCTTCTCTCCGA 660
|||||
QY 1625 CTCCATCTGCTGGCTTCGGTATCAACGCTGAAAAACAACACAGATCTTCTCTTGA 1681
|||||
Db 661 CTCCATCTGCTGGCTTCGGTATCAACGCTG-AAACACACACAGATCTTCTCTTGA 716
|||||
RESULT 5
CD038648 724 bp mRNA linear EST 07-MAY-2003
LOCUS UTTP1010 C11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION Arachis hypogaea cDNA clone UTTP1010_C11 5', mRNA sequence.
ACCESSION CD038648
VERSION 1 GI:30420486
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 724)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
FEATURES
Location/Qualifiers
1..724
source
```

```
/organism="Arachis hypogaea"
/mol_type="mRNA"
/db_xref="taxon:3818"
/clone="UTPPI010_C11"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="X11-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

Query Match 32.3%; Score 657.2; DB 6; Length 724;
Best Local Similarity 95.9%; Pred. No. 1.2e-173;
Matches 697; Conservative 0; Mismatches 24; Indels 6; Gaps 2;

QY 799 CAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACTCAGATCCCATCCGGTTT 858
DB 1 CAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACTCAGATCCCATCCGGTTT 60
QY 859 CATTTCCTACATCTTGAAACCGCCATGACAAACCGAACCCTCAGAGTAGCTAAAATCTCCAT 918
DB 61 CATTTCCTACATCTTGAAACCGCCATGACAAACCGAACCCTCAGAGTAGCTAAAATCTCCAT 120
QY 919 GCCGGTTAAACACCCGGCCAGTTTGAGGATTTCTTCCGGCGAGCGAGACCAATC 978
DB 121 GCCGGTTAAACACCCGGCCAGTTTGAGGATTTCTTCCGGCGAGCGAGACCAATC 180
QY 979 ATCTTACTTCAGGGCTTCAGCAGGATAGCTTGGAGCGCGCTTCAATCGGAATTCAA 1038
DB 181 ATCTTACTTCAGGGATTCAGCAGGATATCTTGGAGCGCGCTTCAATCGGAATTCAA 240
QY 1039 TGAGATACGGAGGTGCTGTTAGAAAGAAATGCAGGAGGTGAGCAAGAGGAGAGGGCA 1098
DB 241 TGAGATACGGAGGTGCTGTTAGAAAGAAATGCAGGAGGTGAGCAAGAGGAGAGGGCA 300
QY 1099 GAGCGATGAGTAGTACTCGGAGTAGTGAGCAATGAAGGAGTAGTAGTCAAGTGCAAA 1158
DB 301 GAGCGCAGGAGTACTCGGAGTAGTG---ATAATGAAGGAGTAGTAGTCAAGTGTCAA 357
QY 1159 GGAGCAGTTGAAGAACTTACTAGCACGCTAAATCCGTCCTCAAGAAAGCGTCCGAAGA 1218
DB 358 GGAGCAGTTCAAGAACTTACTAGCACGCTAAATCCGTCCTCAAGAAAGCGTCC---GA 414
QY 1219 AGAGGAGATATACCAACCAATCAACTTGAGAGAGGGCGAGCCCGCATCTTTCTAACAA 1278
DB 415 AGAGGAGATATACCAACCAATCAACTTGAGAGAGGGCGAGCCCGCATCTTTCTAACAA 474
QY 1279 CTTTGGGAAGTTATTTAGGTGAAGCCAGACAGAGAAACCCCGAGCTTCAGGACCTGGA 1338
DB 475 CTTTGGGAGGTTATTTGAGGTGAAGCCAGACAGAGAAACCCCGAGCTTCAGGACCTGGA 534
QY 1339 CATGATGCTCACCTGTCAGAGATCAAGAGGAGCTTTGATGCTCCACACTTCAACTC 1398
|||||

ORIGIN

535 CATGATGCTCACCTGTCAGAGATCAAGAGAGGAGCTTTGATGCTCCACACTTCAACTC 594
1399 AAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGAACCTTGAACCTCGTGGCTGT 1458
|||||
595 AAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGAACCTTGAACCTCGTGGCTGT 654
1459 AAGAAAGAGCAACACAGAGGGGCGCGGGAAGAGAGGAGGAGGAGCAACGAGAGAGA 1518
|||||
655 AAGAAAGAGCAACACAGAGGGGCGCGGGAAGAGAGGAGTGGGAAGAGAGGAGAGA 714
1519 GGAGGGA 1525
|||||
715 TGAAGA 721
|||||

RESULT 6
CD038172
LOCUS 719 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI003 D10 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI003_D10 5', mRNA sequence.
CD038172
ACCESSION CD038172
VERSION CD038172.1 GI:30420010
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 719)
AUTHORS Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
TITLE Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL Unpublished (2003)
COMMENT Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3
Location/Qualifiers
1. .719
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI003_D10"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="X11-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN	1 (bases 1 to 676)	REFERENCE	1 (bases 1 to 676)
Query Match	31.0%; Score 629.2; DB 6; Length 719;	AUTHORS	Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G. and Lynch,R.E.
Best Local Similarity	95.8%; Pred.No.9.6e-166;	TITLE	Generation and Analyses of ESTs for Arachis hypogaea
Matches 693; Conservative 0; Mismatches 16; Indels 16; Gaps 4;		JOURNAL	Unpublished (2003)
		COMMENT	Contact: Baozhu Guo
QY 1040	GAGATACGGAGGCTGCTGTTAGAGAGAATGCGAGGAGTGAGCAAGAGGAGAGAGGGCGAG 1099		Molecular Genetics
DB 1	GAGATACGGAGGCTGCTGTTAGAGAGAATGCGAGGAGGAGCAAGAGGAGAGAGGGCGAG 60		USDA/ARS, Crop Protection and Management Research Unit
QY 1100	AGCGGATGGAGTACTCGGAGTAGTAGACAATGAAGGAGTGATAGTCAAAGTGTCAAAG 1159		2747 Davis Rd., Tifton, GA 31794, USA
DB 61	AGCGGATGGAGTACTCGGAGTAGTAGACAATGAAGGAGTGATAGTCAAAGTGTCAAAG 117		Tel: 229-387-2334
QY 1160	GACACGTTGAAGAACTTACTAGACGCTTAATCCGTCTCAAGAAAGGCTCCGAAGAA 1219		Fax: 229-387-2321
DB 118	GACACGTTGAAGAACTTACTAGACGCTTAATCCGTCTCAAGAAAGGCTCC---GAA 174		Email: bguo@tifton.usda.gov
QY 1220	GAGGAGATATCAACCAACCAATCAACTTGGAGAGAAGCGGAGCCGATCTTCTAAACAAC 1279		Seq primer: T3.
DB 175	GAGGAGATATCAACCAACCAATCAACTTGGAGAGATGGCGAGCCGATCTTCTAAACAAC 234	FEATURES	Location/Qualifiers
QY 1280	TTTGGGAAGTATTGAGGTGAAGCCAGACAGAAAGAAACCCCGAGCTTCAGACCTGGAC 1339	source	1..676
DB 235	TTTGGGAAGTATTGAGGTGAAGCCAGACAGAAAGAAACCCCGAGCTTCAGACCTGGAC 294		/organism="Arachis hypogaea"
QY 1340	ATGATGCTCACTGCTGAGAGATCAAGAGGAGCTTTGATGCTGCCACACTTCAACTCA 1399		/mol_type="mRNA"
DB 295	ATGATGCTCACTGCTGAGAGATCAAGAGGAGCTTTGATGCTGCCACACTTCAACTCA 354		/cultivar="A13"
QY 1400	AAGGCCATGTTATCGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1459		/db_xref="taxon:3818"
DB 355	AAGGCCATGTTATCGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 414		/clone="UTPPI012_C06"
QY 1460	AGAAAGAGCAACACAGAGGGGACCGCGGAA-----GAAGAGGAGCAGAGAC 1510		/issue_type="Immature pods"
DB 415	AGAAAGAGCAACACAGAGGGGACCGCGGAA-----GAAGAGGAGCAGAGAT 474		/dev_stage="R6"
QY 1511	GAAGAGAGGAGGAGTAAACAGAGAGTGCCTAGGTACACAGCAGAGTTGAAGGAAGGC 1570		/lab_host="XLI-blue"
DB 475	GAAGAGAGGAGGAGTAAACAGAGAGTGCCTAGGTACACAGCAGAGTTGAAGGAAGGC 534		/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
QY 1571	GATGTTGTTATGTCAGCAGCTCATCCAGTACGCCATCAACGCTTCTCCGAACTCCAT 1630		/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
DB 535	GATGTTGTTATGTCAGCAGCTCATCCAGTACGCCATCAACGCTTCTCCGAACTCCAT 594		cDNA library was constructed from peanut cultivar A13;
QY 1631	CTGCTTGGCTTCGGTATCAGCTGAAACACACAGATCTTCC-TTGCAGGTGATAA 1689		(NCV11XAR4). A13 has resistance to Aspergillus infection
DB 595	CTGCTTGGCTTCGGTATCAGCTGAAACACACAGATCTTCTTTCGAGGTGATAA 654		and drought tolerance. The immature pods that developed to
QY 1690	GGACAATGTATAGACCATAGAGAGCAAGCAAGGAGTTTATAGCATTTCCCTGGTCCGG 1749		R6 stage were collected from different plants, and placed
DB 715	TGAAC 1754		into liquid N2 immediately and stored in -80oC freezer.
DB 715	TGAAC 719		Total RNA was isolated with TRIzol-Reagent
RESULT 7			ultrapure(GIBCOBRL). mRNA was extracted and purified from
LOCUS CD038790			total RNA (Promega). cDNA synthesis and library
DEFINITION	UTPPI012 C06 USDA-Tifton Peanut Immature pod cDNA library (UTPP)		construction followed the protocol of by ZAP-cDNA Gigapack
ACCESSION	CD038790		III Gold cloning kit (Stratagene). The cDNA above 500bp
VERSION	CD038790.1 GI:30420628		were collected after size-fraction. The inserts were
KEYWORDS	EST.		directionally cloned into Uni-ZAP XR vector using XhoI
SOURCE	Arachis hypogaea (peanut)		EcoRI sites adapters. The lambda library was packed into
ORGANISM	Arachis hypogaea		phages using Gigapack III Gold (Stratagene). The
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		un-amplified library was used to excise plasmids
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		phagemids from the Uni-ZAP XR vector, and the phagemids
	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;		was used to transform the host bacteria SOLR. The library
	Aeschynomeneae; Arachis.		was constructed by Dr. Meng Luo and Dr. Phat Dang."
		ORIGIN	
		Query Match	30.8%; Score 626; DB 6; Length 676;
		Best Local Similarity	99.0%; Pred.No.7.5e-165;
		Matches 670; Conservative 0; Mismatches 3; Indels 4; Gaps 4;	
		QY 834	ATGCATCAGAAATCCCATCCGGTTTCATTTCTTCTATCATCTTGAACCGCCATGACAAACGAG 893
		DB 1	ATGCATCAGAAATCCCATCCGGTTTCATTTCTTCTATCATCTTGAACCGCCATGACAAACGAG 60
		QY 894	ACCTCAGAGTAGTAAAATCTCCATCCCGTTTAAACACACCCCGCCAGTTTGAAGGATTTCT 953
		DB 61	ACCTCAGAGTAGTAAAATCTCCATCCCGTTTAAACACACCCCGCCAGTTTGAAGGATTTCT 120
		QY 954	TCCCGCGGAGCGGAGACCAATCATCTTCTTCCGGGCTTCAGCGGATACGTTTGG 1013
		DB 121	TCCCGCGGAGCGGAGACCAATCATCTTCTTCCGGGCTTCAGCGGATACGTTTGG 180
		QY 1014	AGCGCGCTTCAATCGGAATTCATAGATACGGAGGTGCTGTTTAGAAGAGAATGCAG 1073
		DB 181	AGCGCGCTTCAATCGGAATTCATAGATACGGAGGTGCTGTTTAGAAGAGAATGCAG 240
		QY 1074	GAGGTGAGCAAGAGGAGAGGGCGAGCGATGAGGATCTCGGAGTAGTGAGAAACAATG 1133
		DB 241	GAGGTGAGCAAGAGGAGAGGGCGAGCGATGAGGATCTCGGAGTAGTGAGAAACAATG 300
		QY 1134	AAGGAGTAGTCTCAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGTAAAT 1193
		DB 301	AAGGAGTAGTCTCAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGTAAAT 360

QY 1194 CCGTCTCAAGAGGCTCCGAAGAGGGAGATATACCAACCCCAATCACTTGAGAG 1253
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 361 CCGTCTCAAGAGGCTCCGAAGAGGGAGATATACCAACCCCAATCACTTGAGAG 420
 QY 1254 AAGCGAGCGCCGATCTTTTAAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGCAAG- 1312
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 421 AAGCGAGCGCCGATCTTTTAAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGCAAGN 480
 QY 1313 AAGAACCCCAAGCTTACGACCTGGAATGATGCTCA-CTGTGTAGAGATCAAGAAGG 1371
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 481 AAGAACCCCAAGCTTACGACCTGGAATGATGCTCACTACCTGTGTAGAGATCAAGAAGG 540
 QY 1372 AGCTTTGATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAGG 1431
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 541 AGCTTTGATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAGG 600
 QY 1432 AACTGGAAACCTTGAACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGG- 1490
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 601 AACTGGAAACCTTGAACTCGTGGCTGT-AGNAAAGAGCAACAACNAGAGGGGACGGCGGGA 659
 QY 1491 AAGAGAGGAGGAGCA 1507
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 660 AAGAGAGGAGGAGCA 676

RESULT 8

CD038555
 LOCUS
 DEFINITION UTPI009 A05 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
 Arachis hypogaea cDNA clone UTPI009_A05 5', mRNA sequence.

ACCESSION CD038555
 SOURCE CD038555.1 GI:30420393
 KEYWORDS EST.

ORGANISM
 Arachis hypogaea (peanut)
 Arachis hypogaea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
 Aeschynomeneae; Arachis.
 REFERENCE 1 (bases 1 to 685)
 AUTHORS Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
 and Lynch, R. E.

TITLE Generation and Analyses of ESTs for Arachis hypogaea
 JOURNAL Unpublished (2003)
 COMMENT Contact: Baozhu Guo
 Molecular Genetics
 USDA/ARS, Crop Protection and Management Research Unit
 2747 Davis Rd, Tifton, GA 31794, USA
 Tel: 229-387-2334
 Fax: 229-387-2321
 Email: bguo@tifton.usda.gov
 Seq primer: T3.

FEATURES

source
 1..685
 /organism="Arachis hypogaea"
 /mol_type="mRNA"
 /cultivar="A13"
 /db_xref="taxon:3818"
 /clone="UTPI009 A05"
 /tissue_type="Immature pods"
 /dev_stage="R6"
 /lab_host="X11-blue"
 /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
 /note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
 cDNA library was constructed from peanut cultivar A13
 (NCV1XAR). A13 has resistance to Aspergillus infection
 and drought tolerance. The immature pods that developed to
 R6 stage were collected from different plants, and placed
 into liquid N2 immediately and stored in -80oC freezer.
 Total RNA was isolated with TRIzol-Reagent
 ultrapure (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of By ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 29.6%; Score 602; DB 6; Length 685;
 Best Local Similarity 94.3%; Pred. No. 4.5e-158;
 Matches 663; Conservative 0; Mismatches 21; Indels 19; Gaps 3;
 QY 29 ATATAAGTAGTAGCAGAGCAATGAGAGGGAGGTTTCTCCACTGATGCTGTGCTAGGG 88
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATATAAGTAGTAGCAGAGCAATGAGAGGGAGGTTTCTCCACTGATGCTGTGCTAGGG 60
 QY 89 ATCTTTGTCTGGCTTTCAAGTCATGCCAAGCATGCCAAGTCATCACCTTACCAAGAA 148
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 ATCTTTGTCTGGCTTTCAAGTCATGCCAAGTCATCACCTTACCAAGAA 114
 QY 149 ACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTG 208
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 115 ACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGACGACTTG 174
 QY 209 AAGCAAAAGCATCGAGTCTCGCTCCACCAAGCTCGAGTATGATCTCTGCTGTGCTAT 268
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 175 AAGCAAAAGCATCGAGTCTCGCTGCACCAAGCTCGAGTATGATCTCTGCTGTGCTAT 234
 QY 269 GATCTCTGAGGACACACTGCGCACCAACCAACAGTTCCTCCAGGGAGCGGACACGT 328
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 235 G-----ACACTGGCGCCACCAACCAACGTCACCTCCAGGGAGCGGACACGT 282
 QY 329 GCGCGCAACCCGGAGACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 283 GCGCGCAACCCGGAGACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 QY 389 CGATGGGACCACTGCGACCGGAGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 343 CGATGGGACCACTGCGACCGGAGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
 QY 449 GAAGATTGGAGCGCACCAAGTCACTAGCAGCCACCGAAATAAGGCCCGGAGGAGAGAGAA 508
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 403 GAAGATTGGAGCGCACCAAGTCACTAGCAGCCACCGAAATAAGGCCCGGAGGAGAGAGAA 462
 QY 509 GGAGAACAGAGTGGGGAACACAGGTAGCATGTGAGGGAAGAAACATCTC-GGAACAA 567
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 463 GGAGAACAGAGTGGGGAACACAGGTAGCATGTGAGGGAAGAAACATCACCGGAACAA 522
 QY 568 CCCTTTCTACTTCCCTCAAGGCGGTTAGCACCCCTAGCGGGAACCAACCGTAGGAT 627
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 523 CCCTTTCTACTTCCCTCAAGGCGGTTAGCACCCCTAGCGGGAACCAACCGTAGGAT 582
 QY 628 CCGGGTCTCGAGAGGTTTGACCAAGGTCAAGCGAGTTTCAAGATCTCCAGAAATCACCG 687
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 583 CCGGTCTCGAGAGGTTTGACCAAGGTCAAGCGAGTTTCAAGATCTCCAGAAATCACCG 642
 QY 688 TATTGCGAGTGGGCGCAACCTAACACTCTTGTGTTCTTCCC 730
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 643 TATTGCGAGTGGGCGCAGACCTAACACTCTTGTGTTCTTCCC 685

RESULT 9

CD038119
 LOCUS
 DEFINITION UTPI002 F11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
 Arachis hypogaea cDNA clone UTPI002_F11 5', mRNA sequence.
 ACCESSION CD038119
 VERSION CD038119.1 GI:30419957

```

KEYWORDS
SOURCE
ORGANISM

EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 588)
AUTHORS
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
CONTACT: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
FEATURES
source
1..588
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI002.F11"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCv11XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."
ORIGIN
Query Match 28.3%; Score 576; DB 6; Length 588;
Best Local Similarity 99.8%; Pred. No. 9.3e-151;
Matches 587; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1438 AAACCTTGAACTCGTGGCTGTAGAAAGAGGCAACACAGAGGGGACGGGGAAGAAGA 1497
DB 1 AAACCTTGAACTCGTGGCTGTAGAAAGAGGCAACACAGAGGGGACGGGGAAGAAGA 60
QY 1498 GGAGGACGAAGACGAAGAGAGGAGGAGTAAACAGAGAGGTCGCTAGGTACACAGCGAG 1557
DB 61 GGNAGACGAAGACGAAGAGAGGAGGAGGAGTAAACAGAGAGGTCGCTAGGTACACAGCGAG 120
QY 1558 GTTGAAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTC 1617
DB 121 GTTGAAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTC 180
QY 1618 CTCGGAATCCATCTGCTGGCTTCGGTATCAACGCTGAAACACACAGATCTTCTCT 1677
DB 181 CTCGGAATCCATCTGCTGGCTTCGGTATCAACGCTGAAACACACAGATCTTCTCT 240

```

```

QY 1678 TGCAGGTGATAAGGACAAATGTGATAGACACAGATAGAGCAAGCAAGGATTTAGCAATT 1737
DB 241 TCAGGTGATAGGACAAATGTGATAGACACAGATAGAGCAAGCAAGGATTTAGCAATT 300
QY 1738 CCCTGGTGGGTGGAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATTCACCTTTGT 1797
DB 301 CCCTGGTGGGTGGAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATTCACCTTTGT 360
QY 1798 GAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 1857
DB 361 GAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 420
QY 1858 GAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1916
DB 421 GAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 1917 AGGCTTTTAACTGAGAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1976
DB 481 AGGCTTTTAACTGAGAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 1977 TACTCTACTATCCAAAACTTATCAATAATAAAAAAGCTTTTGTGCGTT 2024
DB 541 TACTCTACTATCCAAAACTTATCAATAATAAAAAAGCTTTTGTGCGTT 588

RESULT 10
LOCUS
DEFINITION
CD038540 684 bp mRNA linear EST 07-MAY-2003
Arachis hypogaea cDNA clone UTPPI008_G09 5', mRNA sequence.
ACCESSION
CD038540
VERSION
CD038540.1 GI:30420378
KEYWORDS
EST.
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 684)
AUTHORS
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
CONTACT: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
FEATURES
source
1..684
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI008_G09"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCv11XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

```

construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

```
Query Match          26.0%; Score 527.6; DB 6; Length 684;
Best Local Similarity 94.3%; Pred. No. 4.5e-137; Indels 15; Gaps 3;
Matches 584; Conservative 0; Mismatches 20;

Qy 957 CGGCGAGCAGCGGAGACCAATCATCTTACCTTGCAGGGCTTCAGCAGGAATACGTTGGAGG 1016
Db 72 CGGCGAGCAGCGGAGACCAATCATCTTACCTTGCAGGGATTCAGCAGGAATACTTTGGAGG 131
Qy 1017 CCGCTTCAATGCGGAATTCATGAGATACGGAGGGTCTGTTAGAGAGAGATCCAGGAG 1076
Db 132 CCGCTTCAATGCGGAATTCATGAGATACGGAGGGTCTGTTAGAGAGAGATCCAGGAG 191
Qy 1077 GTGAGCAAGAGGAGAGAGGGCAGAGCGATCGAGTACTCGGAGTAGTGAGAACATGAAG 1136
Db 192 GAGAGCAAGAGGAGAGAGGGCAGAGCGGAGTACTCGGAGTAGTG---ATAATGAAG 248
Qy 1137 GAGTGATAGTCAAAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTAAGCAGCCTAAATCCG 1196
Db 249 GAGTGATAGTCAAAAGTGTCAAAGGAGCAGCGTTCAAGAACTTACTAAGCAGCCTAAATCCG 308
Qy 1197 TCTCAAGAAAGGCTCCGAGAGAGAGGAGATATCAACACCCCAATCAACTTGAAGAAAG 1256
Db 309 TCTCAAGAAAGGCTCC---GAAGAGGAAGATATCAACACCCCAATCAACTTGAAGAGATG 365
Qy 1257 GCGAGCCGAGTCTTTCTAACAACTTTGGAGGTATTATTAGGTCAAGCCAGCAGAGAAGA 1316
Db 366 GCGAGCCGAGTCTTTCTAACAACTTTGGAGGTATTATTAGGTCAAGCCAGCAGAGAAGA 425
Qy 1317 ACCCCAGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAGAGAGGAGCTT 1376
Db 426 ACCCCAGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAGAGAGGAGCTT 485
Qy 1377 TGATGCTCCACACTTCAACTCAAAAGGCCATGTTATGCTGCTCAACAAAGGAAGACTG 1436
Db 486 TGATGCTCCACACTTCAACTCAAAAGGCCATGTTATGCTGCTCAACAAAGGAAGACTG 545
Qy 1437 GAAACCTTGAACCTGCTGTAGAAAGAGCAACACAGAGGGGACGGCGAA--- 1492
Db 546 GAAACCTTGAACCTGCTGTAGAAAGAGCAACACAGAGGGGACGGCGGAAACAAG 605
Qy 1493 -----GAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGTGCGTAGGT 1547
Db 606 AGTGGGAAGAGAGAGGAGAGATGAAGAGAGGAGGAGGTAAACAGAGAGTGCGTAGGT 665
Qy 1548 ACACAGGAGGTTCAAGGA 1566
Db 666 ACACAGGAGGTTGANGAA 684
```

RESULT 11

```
CD038694
LOCUS          CD038694          474 bp      mRNA      linear      EST 07-MAY-2003
DEFINITION    UTPII010_H08 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
               Arachis hypogaea cDNA clone UTPII010_H08 5', mRNA sequence.
ACCESSION     CD038694
VERSION       CD038694.1 GI:30420532
KEYWORDS      EST.
SOURCE        Arachis hypogaea (peanut)
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

roside; eurouside I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
1 (Bases 1 to 474)
Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G. and Lynch, R. E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: F3.

FEATURES
source

```
1. .474
/Location/Qualifiers
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPII010_H08"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

ORIGIN

```
Query Match          23.3%; Score 473; DB 6; Length 474;
Best Local Similarity 99.8%; Pred. No. 1e-121; Indels 0; Gaps 0;
Matches 473; Conservative 0; Mismatches 1;

Qy 1559 TTGAAGAAGCGGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAAGCTTCC 1618
Db 1 TTGAAGAAGCGGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAAGCTTCC 60
Qy 1619 TCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACCAACAGAACTTCCTT 1678
Db 61 TCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACCAACAGAACTTCCTT 120
Qy 1679 GCAGGTGATAGGACATGTAGTACCAGATAGAGAGCAAGCGAGGATTTAGCATTC 1738
Db 121 GCAGGTGATAGGACATGTAGTACCAGATAGAGAGCAAGCGAGGATTTAGCATTC 180
Qy 1739 CTTGGGTTCGGGTGAAACAAGTTGAGAAGCTCATCAAAAACCAAGAAAGTAATCTCACTTTGTG 1798
Db 181 CTTGGGTTCGGGTGAAACAAGTTGAGAAGCTCATCAAAAACCAAGAAAGTAATCTCACTTTGTG 240
Qy 1799 AGTGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAG 1858
Db 241 AGTGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAG 300
Qy 1859 AAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
```

```

|||||
301 AAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCTCTTCAATTTGAAG 360
|||||
1919 GCTTTTAACCTAGATGGAGCAACTTGTATGATCGATAATAAGATCAAGCTTTTGTGA 1978
|||||
361 GCTTTTAACCTAGATGGAGCAACTTGTATGATCGATAATAAGATCAAGCTTTTGTGA 420
|||||
1979 CTCTACTATCAAAAACCTTATCAATAATAAAAACGTTTGTGGTGTCTTCTCC 2032
|||||
421 CTCTACTATCAAAAACCTTATCAATAATAAAAACGTTTGTGGTGTCTTCTCC 474
|||||

RESULT 12
LOCUS CD038620
DEFINITION UTPI009_H07 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION CD038620
VERSION CD038620.1 GI:30420458
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 582)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2334
Email: bguo@tifton.usda.gov
Seq primer: T3
Location/Qualifiers
1..582
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPI009_H07"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NC111X44). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were directionally cloned after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

FEATURES
source
1..582
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPI009_H07"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NC111X44). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were directionally cloned after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match 23.2%; Score 471; DB 6; Length 582;

```

```

Best Local Similarity 91.5%; Pred. No. 4e-121;
Matches 552; Conservative 0; Mismatches 19; Indels 32; Gaps 4;

QY 1398 CAAAGCCCATGTTATCTCGTCGTCGTCACAAAGGAACTGGAACCTTGAACCTGCTGGCTG 1457
|||||
DB 1 CAAAGCCCATGTTATCTCGTCGTCGTCACAAAGGAACTGGAACCTTGAACCTGCTGGCTG 60
|||||
QY 1458 TAAGAAAAGAGCAACAACAGAGGGGACGGCGGAA-----GAAGAGGAGGAGCAAG 1508
|||||
DB 61 TAAGAAAAGAGCAACAACAGAGGGGACGGCGGAAAGAGAGGAGGAGGAAG 120
|||||
QY 1509 ACAGAAAGAGAGGAGGAGTAACAGAGAGGTCGTAAGTACACAGAGAGGTTGAAGGAAG 1568
|||||
DB 121 ATGAAGAGAGAGGAGGAGTAACAGAGAGGTCGTAAGTACACAGAGAGGTTGAAGGAAG 180
|||||
QY 1569 GCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCC 1628
|||||
DB 181 GCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCC 240
|||||
QY 1629 ATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTGCAGGTGATA 1688
|||||
DB 241 ATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTGCAGGTGATA 300
|||||
QY 1689 AGGACAAATGTATAGACCAAGATAGAGAACCAAGCGAAGGATTAGCATTTCCCTGGGTGCG 1748
|||||
DB 301 AGGACAAATGTATAGACCAAGATAGAGAACCAAGCGAAGGATTAGCATTTCCCTGGGTGCG 360
|||||
QY 1749 GTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAGTCTCGTC 1808
|||||
DB 361 GTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAGTCTCACTTTGCGAGTCTCGTC 420
|||||
QY 1809 CTCATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATC 1868
|||||
DB 421 CTCATCTCAATCTCGTC-----GTCTCTGAGAAAGAGGATC 459
|||||
QY 1869 AAGAGAGAGAAAAACCAAGGA-GGGAGGCTCACTCTCTTTCATTTTGAAGGCTTTTAAAC 1927
|||||
DB 460 AAGAGAGAGAAAAACCAAGGAGGGAAGGTCCTCACTCTTTTCAATTTTGAAGGCTTTTAAAC 519
|||||
QY 1928 TCAGATGAGGAGCAACTTGTATGTAT-CGATAAAGATCACGCTTTTGTACTCTACTA 1986
|||||
DB 520 TGAGAAATGAGGAANCTTGTATGTATCCATTAATAAGATCACGCTTTTGTACTCTACTA 579
|||||
QY 1987 TCC 1989
DB 580 TCC 582

RESULT 13
LOCUS CD038253
DEFINITION UTPI004_F03 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION CD038253
VERSION CD038253.1 GI:30420091
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 509)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2334

```

[illegible]

```
Db      62  GAAGCGGATGTGTTTCATCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAA 121
QY      1625 CTCATCTGCTTGGCTTCGGTATCAACGCTGAACCAACACACAGAATCTTCTTGCAAGT 1684
Db      122 CTCATCTGCTTGGCTTCGGTATCAACGCTGAACCAACACACAGAATCTTCTTGCAAGT 181
QY      1685 GATAAGGACAATGTGATAGACCAAGTAGAGAACCAAGCGAAGATTTAGCATTTCCCTGGG 1744
Db      182 GATAAGGACAATGTGATAGACCAAGTAGAGAACCAAGCGAAGATTTAGCATTTCCCTGGT 241
QY      1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGAATCTCACTTTGTGAGTGTCT 1804
Db      242 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGGAGTCTCACITTTGTGAGTGTCT 301
QY      1805 CGTCCTCAATCTCAATCTCAATCTCGTCTCCTCGAGAAAGAGTCTCCTGAGAAAGAG 1864
Db      302 CGTCCTCAATCTCAATCTCGGTC-----GTCCTCTGAAAAAGAG 340
QY      1865 GATCAAGAGGAGGAAAAACCAAGGAGGAGGGTCCACTCCTTTCAATTTTGAAGGCTTTT 1924
Db      341 GATCAAGAGGAGGAAAAACCAAGGAGGAGGGTCCACTCCTTTCAATTTTGAAGGCTTTT 400
QY      1925 AACTGAGAAATGGAGGAACCTGTTATGTATCGATAATAGATCACGCTTTTGTACTCTAC 1984
Db      401 AACTGAGAAATGGAGGAACCTGTTATGTATCCATAATAAGATCACGCTTTTGTAACTTAC 460
QY      1985 TATCCAAAACCTTATCAATAAATA 2008
Db      461 TATCCAAAACCTTATCAATAAATA 484
```

RESULT 15

```
CO897502
LOCUS      CO897502
DEFINITION EST00007 Peanut Lambda Express library Arachis hypogaea cDNA 5',
            mRNA sequence.
ACCESSION CO897502
VERSION   CO897502.1 GI:51237292
KEYWORDS  EST.
SOURCE    Arachis hypogaea (peanut)
           ORGANISM
           Arachis hypogaea
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
           Aeschynomeneae; Arachis.
           1 (bases 1 to 443)
           Yan, Y.S., Wang, L. and Huang, S.Z.
           Expressed sequence tags of mid-matured cotyledons of peanut
           Unpublished (2004)
           Contact: ShangZhi Huang
           Plant Development and Molecular Biology
           Sun Yat-sen University, Department of Biology and Biotechnology
           Guangzhou, 510275, China
           Tel: (860) 02084036592
           Email: yys9803@yahoo.com.cn
           Seq primer: pTriplex2 Forward.
           Location/Qualifiers
             1..443
               /organism="Arachis hypogaea"
               /mol_type="mRNA"
               /cultivar="Yueyou 523"
               /db_xref="taxon:3818"
               /tissue_type="Cotyledons"
               /dev_stage="Mid-matured stage"
               /lab_host="E.coli BM25.8"
               /clone_lib="Peanut Lambda Express library"
               /note="Organ: Seed; Vector: lambdaTriplex2"
```

ORIGIN

```
Query Match      21.1%; Score 428.6; DB 7; Length 443;
Best Local Similarity 98.0%; Pred. No. 3.4e-109;
Matches 434; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      597 GCACCGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAAGT 656
Db      1 GCACCGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAAGT 60
QY      657 CAAAGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCGAGATCGAGGCCAAACCTTAAAC 716
Db      61 CAAAGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCGAGATCGAGGCCAGACCTTAAAC 120
QY      717 CTCCTTGTCTTCCCAAGCAGCGCTGATGCTGATAACATCCTTGTATCCAGCAGGCGCAAG 776
Db      121 CTCCTTGTCTTCCCAAGCAGCGCTGATGCTGATAACATCCTTGTATCCAGCAGGCGCAAG 180
QY      777 CCACCGTGACCGTAGCAAAATGCAATAACAGAAAAGAGCTTTTAATCTTGAACGAGGCCATG 836
Db      181 CCACCGTGACCGTAGCAAAATGCAATAACAGAAAAGAGCTTTTAATCTTGAACGAGGCCATG 240
QY      837 CACTCAGAAATCCATCCGGTTTCATCTTCTACATCTTGAACCGCCATGACCAACCAAGCC 896
Db      241 CACTCAGAAATCCATCCGGTTTCATCTTCTACATCTTGAATCGACATGACCAACCAAGCC 300
QY      897 TCAGAGTAGCTAAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTTGAGGATTTCTTCC 956
Db      301 TCAGAGTAGCTAAAATCTCCATGCCGTTTAAACACCCGCGCAGTTTGAGGATTTCTTCC 360
QY      957 CGCGCAGCAGCGAGACCAATCATCTACTTGCAGGGCTTTCAGCAGGAATACGTTTGGAGG 1016
Db      361 CGCGCAGCAGCGAGACCAATCATCTACTTGCAGGGATTCAGCAGGAATACGTTTGGAGG 420
QY      1017 CGGCCTTCAATCGGGAATTCAAAT 1039
Db      421 CGGCCTTCAATCGGGAATTCAAAT 443
```

Search completed: August 24, 2005, 03:44:08

Job time : 4769.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 760.871 Seconds
(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 2032

Sequence: 1 aataatcatatattatcatc.....cgtttgctggttggtttttccc 2032

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	4 AAF90339	Aaf90339 Peanut al
2	2032	100.0	2032	8 ABX70603	Abx70603 Peanut Ar
3	2032	100.0	2032	10 ADG27462	Adg27462 Peanut Ar
4	2032	100.0	2041	4 AAS08537	Aas08537 DNA encod
5	2025.6	99.7	2032	2 AAT76613	Aat76613 Peanut al
6	1930	95.0	1930	2 AAZ06382	Aaz06382 Peanut al
7	1752.4	86.2	1949	2 AAT76612	Aat76612 Peanut al
8	1752.4	86.2	1949	8 ABX70604	Abx70604 Peanut Ar
9	1750.8	86.2	1949	10 ADG27463	Adg27463 Peanut Ar
10	1749.8	86.1	1952	4 AAS08539	Aas08539 Anaphylac
11	408.4	20.1	1251	6 ABS55196	Abs55196 Glycine m
12	406.8	20.0	1254	9 ACC49562	Acc49562 Mature be
13	406.8	20.0	1278	9 ACC49561	Acc49561 FLAG-tag
14	406.8	20.0	1320	2 AAV17564	Aav17564 Coding se
15	406.8	20.0	1320	2 ACC49553	Acc49553 Glycine m
16	404	19.9	1632	6 ABS55198	Abs55198 Glycine m
17	404	19.9	1818	2 AAV17562	Aav17562 Coding se
18	402	19.8	1680	6 ABS55197	Abs55197 Glycine m
19	324.2	16.0	1920	2 AAV17563	Aav17563 Coding se
20	160.2	7.9	2140	2 AAV42316	Aav42316 Macadamia

21	157.4	7.7	2171	2 AAV42311	Aav42311 Macadamia
22	152.6	7.5	2171	2 AAV42310	Aav42310 Macadamia
23	138.8	6.8	1867	2 AAQ20377	Aaq20377 Sequence
24	98.8	4.9	1924	2 AAV72243	Aav72243 G. max SB
25	88.4	4.4	1580	12 ADE80950	Ade80950 Cashew nu
26	88.4	4.4	1730	12 ADE80949	Ade80949 Cashew nu
27	79.6	3.9	584	13 ACN47495	Acn47495 Cotton pr
28	79.4	3.9	2332	13 ADR63468	Adr63468 Cotton CD
29	74	3.6	1722	10 ADL18492	Adl18492 Maize glo
30	74	3.6	2003	10 ADL18494	Adl18494 Maize glo
31	66	3.2	601	13 ACN47683	Acn47683 Cotton pr
32	59.4	2.9	635	13 ACN46465	Acn46465 Cotton pr
33	57.8	2.8	2000	8 ADA71938	Ada71938 Rice gene
34	56.6	2.8	574	13 ACN46319	Acn46319 Cotton pr
35	53	2.6	2010	10 ADC08559	Adc08559 Rice DNA
36	51.2	2.5	659	13 ACN50255	Acn50255 Cotton no
37	50.2	2.5	193853	11 ACN4956	Acn4956 Mouse gen
38	49.6	2.4	10732	3 AAA10594	Aaa10594 Gene enco
39	49.4	2.4	1068	4 AAD05672	Aad05672 Human zma
40	49.4	2.4	1068	10 ADB49430	Adb49430 Degenerat
41	48.8	2.4	640	13 ADR63469	Adr63469 Cotton CD
42	47.6	2.3	591	13 ACN50453	Acn50453 Cotton ma
43	46.6	2.3	559	13 ACN48185	Acn48185 Cotton pr
44	46.6	2.3	90401	12 ADQ97515	Adq97515 Human can
45	46	2.3	564	13 ACN49247	Acn49247 Cotton pr

ALIGNMENTS

RESULT 1
AAF90339
ID AAF90339 standard; cDNA; 2032 BP.
XX
AC AAF90339;
XX
DT 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX
DE Peanut allergen Ara hl P41B cDNA.
XX
KW Peanut; allergen; Ara hl P41B; transgenic plant; allergy; ss.
XX
OS Arachis hypogaea.
XX
FH Key Location/Qualifiers
FT misc_feature 676..930
FT /*tag= a
FT /*note= "PCR amplified fragment"
XX
FN WO200136621-A2.
XX
PD 25-MAY-2001.
XX
PF 20-NOV-2000; 2000WO-US031657.
XX
PR 19-NOV-1999; 99US-0167255P.
XX
PA (UYAL-) UNIV ALABAMA A & M.
XX
PI Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX
WPI; 2001-355630/37.
XX
DR Producing transgenic peanut plants that produce allergen-free seeds,
XX
PT useful in non-allergenic foods, by antisense or sense co-suppression of
XX
PT allergen-encoding genes.
XX
PS Claim 20; Fig 5; 72pp; English.
XX
CC The present sequence is that of peanut allergen Ara hl P41B cDNA. A
XX
CC portion of this gene is homologous to the corresponding region of the
XX
CC peanut allergen Ara hl P17 gene. This region has been PCR amplified,

CC cloned in transformation vectors (pUC18 and pBI4434) in sense and
CC antisense orientations and used to down-regulate Ara h1 p41b and Ara h1
CC p17 allergens in peanut. This is an example of the method of the
CC invention, which relates to the production of a peanut plant having
CC reduced, or undetectable, allergenic protein (AP) content in its seed. A
CC peanut plant cell is transformed with a DNA construct containing an
CC antisense AP gene and/or sense AP gene, or their fragments, regenerated
CC to plants, and fertile transgenic plants that produce seeds with reduced
CC AP content are identified. The seeds are useful for preparation of
CC allergen-free foods. (Updated on 11-SEP-2003 to standardise OS field)
XX

SQ Sequence 2032 BP; 628 A; 473 C; 530 G; 401 T; 0 U; 0 Other;

Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAG 60
DB |||||
QY 61 GGTTCCTCACATGCTGCTAGGATCCTTCCTGGCTTCAGTTCTGCAACGCA 120
DB |||||
QY 121 TGCCAAGTCATCACCTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCTCCA 180
DB |||||
QY 121 TGCCAAGTCATCACCTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCTCCA 180
QY 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCTGCTGCAACAA 240
DB |||||
QY 241 GCTCGAGTATGATCCTCGTTGTTCTATGATCCTCGAGGACACATGCGCACCAACCA 300
DB |||||
QY 301 AGTTTCCCTCCAGGGAGCGACACGTTGGCGCCAAACCCGAGACTACGATGATGACCG 360
DB |||||
QY 301 AGTTTCCCTCCAGGGAGCGACACGTTGGCGCCAAACCCGAGACTACGATGATGACCG 360
QY 361 CGGTCAACCCCGAAGAGAGAAAGGCGGATGCGGGAACAGTGTGACCGAGGAGCGTGA 420
DB |||||
QY 361 CGGTCAACCCCGAAGAGAGAAAGGCGGATGCGGGAACAGTGTGACCGAGGAGCGTGA 420
QY 421 AAGAGAGAGAGACTGGAGACAAACCAAGAGAGATTTGGAGGGGACCAAGTTCATCAGCAGCC 480
DB |||||
QY 481 ACGGAAATTAAGCCCGAAGAGAGAGAGAGAAACAGAGTGGGGAACACCAAGGTAGCCA 540
DB |||||
QY 481 ACGGAAATTAAGCCCGAAGAGAGAGAGAGAGAAACAGAGTGGGGAACACCAAGGTAGCCA 540
QY 541 TGTGAGGAGAGAAACATCTCGGAAACAAACCTTTCTACTTCCGTCGTCGAGCGGTTTAGCAC 600
DB |||||
QY 541 TGTGAGGAGAGAAACATCTCGGAAACAAACCTTTCTACTTCCGTCGTCGAGCGGTTTAGCAC 600
QY 601 CGCTACCGGAGCAACCAAGGTAGGATCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
DB |||||
QY 601 CGCTACCGGAGCAACCAAGGTAGGATCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
QY 661 GCAGTTTCAGATCTCCAGATCACCGTATTTGTCAGATCGAGGCGCAACCTTAACTCT 720
DB |||||
QY 661 GCAGTTTCAGATCTCCAGATCACCGTATTTGTCAGATCGAGGCGCAACCTTAACTCT 720
QY 721 TGTTCTCCCAAGACGCTGATGCTGATAACATCTTGTATTCAGAGAGGCAAGCCAC 780
DB |||||
QY 721 TGTTCTCCCAAGACGCTGATGCTGATAACATCTTGTATTCAGAGAGGCAAGCCAC 780
QY 781 CGTGACCTAGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840
DB |||||
QY 781 CGTGACCTAGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840

QY 841 CAGAAATCCATCCGCTTTCATTTCTCAATCTTGAAACCGCCATGATAACACCAAGAACTCAG 900
DB |||||
QY 841 CAGAAATCCATCCGCTTTCATTTCTCAATCTTGAAACCGCCATGATAACACCAAGAACTCAG 900
QY 901 AGTAGCTAAAATCTCCATGCGCGTTTAAACACACCCGCGCAGTTTGTAGGATTTCTTCCCGGC 960
DB |||||
QY 901 AGTAGCTAAAATCTCCATGCGCGTTTAAACACACCCGCGCAGTTTGTAGGATTTCTTCCCGGC 960
QY 961 GAGCAGCCGAGACCAATCATCTTCTCAGCGGCTTTCAGCAGGAATACGTTGGAGCCGCG 1020
DB |||||
QY 961 GAGCAGCCGAGACCAATCATCTTCTCAGCGGCTTTCAGCAGGAATACGTTGGAGCCGCG 1020
QY 1021 CTTCAATCGGAATTTCAATGAGATACGAGGCTGCTTTAGAAGAGAAATGCAAGAGGTGA 1080
DB |||||
QY 1021 CTTCAATCGGAATTTCAATGAGATACGAGGCTGCTTTAGAAGAGAAATGCAAGAGGTGA 1080
QY 1081 GCAAGAGAGAGAGAGGCGCAGTGGAGTACTCGGAGTAGTAGAGAACATGAAGGAGT 1140
DB |||||
QY 1081 GCAAGAGAGAGAGGCGCAGTGGAGTACTCGGAGTAGTAGAGAACATGAAGGAGT 1140
QY 1141 GATAGTCAAGTGTCAAAAGGAGCAGTTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTCTC 1200
DB |||||
QY 1141 GATAGTCAAGTGTCAAAAGGAGCAGTTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTCTC 1200
QY 1201 AAAGAAAGGCTCCGAAAGAGGAGGAGATATCAACCAACCAATCACTTTGAGAGAGGCGA 1260
DB |||||
QY 1201 AAAGAAAGGCTCCGAAAGAGGAGGAGATATCAACCAACCAATCACTTTGAGAGAGGCGA 1260
QY 1261 GCCCGATCTTTCTAAACCTTTGGGAGTTATTTGAGGTGAAGCCAGACAGAGAAACCC 1320
DB |||||
QY 1261 GCCCGATCTTTCTAAACCTTTGGGAGTTATTTGAGGTGAAGCCAGACAGAGAAACCC 1320
QY 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGGCTTTGAT 1380
DB |||||
QY 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGGCTTTGAT 1380
QY 1381 GCTCCCACTTTCAACTCAAAAGGCCATGTTTATCGTCGTCTCAACAAAGGAACCTGGA 1440
DB |||||
QY 1381 GCTCCCACTTTCAACTCAAAAGGCCATGTTTATCGTCGTCTCAACAAAGGAACCTGGA 1440
QY 1441 CTTTGAATCTGTGCTGTAAAGAAAGAGCAACACAGAGGGGACGCGCGGAGAAAGAGGA 1500
DB |||||
QY 1441 CTTTGAATCTGTGCTGTAAAGAAAGAGCAACACAGAGGGGACGCGCGGAGAAAGAGGA 1500
QY 1501 GAGCAGAACCAAGAGAGGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTT 1560
DB |||||
QY 1501 GAGCAGAACCAAGAGAGGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTT 1560
QY 1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCTC 1620
DB |||||
QY 1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCTC 1620
QY 1621 CGAATCTCATCTGTTGCTTGGCTTCAACGCTGAAACCAACCAAGAACTTCTTCTTGC 1680
DB |||||
QY 1621 CGAATCTCATCTGTTGCTTGGCTTCAACGCTGAAACCAACCAAGAACTTCTTCTTGC 1680
QY 1681 AGGTGATAAGCACAATGTGATAGACCCAGATAGAGCAAGCAAGAGATTTAGCATTTCCC 1740
DB |||||
QY 1681 AGGTGATAAGCACAATGTGATAGACCCAGATAGAGCAAGCAAGAGATTTAGCATTTCCC 1740
QY 1741 TGGGTGCGGTGAACAGATTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTTGTGAG 1800
DB |||||
QY 1741 TGGGTGCGGTGAACAGATTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTTGTGAG 1800
QY 1801 TGCTGCTCCTCAATCTCAATCTCAATCTCCGCTGCTCCTGAGAAAGAGTCTCTTGAGAA 1860
DB |||||
QY 1801 TGCTGCTCCTCAATCTCAATCTCAATCTCCGCTGCTCCTGAGAAAGAGTCTCTTGAGAA 1860
QY 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
DB |||||
QY 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
QY 1921 TTTTAACTGAGAAATGGAGGCAACTTTGTTATGTTATGATAAAGATCAACGCTTTTGTACT 1980

901	AGTAGCTAAAAATCTCCATGCGCCGTTAAACACACCCCGGCAGTTTGAGGAAATTTCTTCCCGGC	960
961	GAGCAGCCGACACCAATCATCTACTTCCAGGGCTTCAGCAGGAATACGTTGGAGGCGGC	1020
961	GAGCAGCCGACACCAATCATCTACTTTCAGGGCTTCAGCAGGAATACGTTGGAGGCGGC	1020
1021	CTTCAATGCGGAATTCAAATAGATACGGAGGGTCTGTTTAGAGAGAAATGCAGGAGTGA	1080
1021	CTTCAATGCGGAATTCAAATAGATACGGAGGGTCTGTTTAGAGAGAAATGCAGGAGTGA	1080
1081	GCAAGAGGAGAGCGGACGAGCGCATCGAGTACTCGGAGTAGTCAGAACCAATGAAGGAGT	1140
1081	GCAAGAGGAGAGCGGACGAGCGCATCGAGTACTCGGAGTAGTCAGAACCAATGAAGGAGT	1140
1141	GATAGTCAAAGTGTCAAAGGAGCAGTTGAAAGAACTTACTAAGCACGCTAAATCCGCTCTC	1200
1141	GATAGTCAAAGTGTCAAAGGAGCAGTTGAAAGAACTTACTAAGCACGCTAAATCCGCTCTC	1200
1201	AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTGTAGAGAAAGCGGA	1260
1201	AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTGTAGAGAAAGCGGA	1260
1261	GCCCGATCTTTCTAAACAATTTTGGAAAGTTATTTTGGGTGAAGCCAGCACAGAGAAACCC	1320
1261	GCCCGATCTTTCTAAACAATTTTGGAAAGTTATTTTGGGTGAAGCCAGCACAGAGAAACCC	1320
1321	CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGAT	1380
1321	CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGAT	1380
1381	GCTCCCACTTTCAACTCAAAAGGCCATGGTTATCGTCGTCAACAAAGAACTCGAAA	1440
1381	GCTCCCACTTTCAACTCAAAAGGCCATGGTTATCGTCGTCAACAAAGAACTCGAAA	1440
1441	CTTTGAACTCGTGGCTGTTAAGAAAGGACCAACACAGAGGGGACGCGCGGAAGAGAGGA	1500
1441	CTTTGAACTCGTGGCTGTTAAGAAAGGACCAACACAGAGGGGACGCGCGGAAGAGAGGA	1500
1501	GGACGAAGACGAAGAAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
1501	GGACGAAGACGAAGAAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
1561	GAAGGAAGGCGATGTTTCACTATGCACGACAGCTCATCCAGTAGCCATCAACCGTTCCCTC	1620
1561	GAAGGAAGGCGATGTTTCACTATGCACGACAGCTCATCCAGTAGCCATCAACCGTTCCCTC	1620
1621	CGAACTCCATCTGCTTCGGCTTCGGTATCAACGCTGAANAACACACAGAAATCTTCCTTGC	1680
1621	CGAACTCCATCTGCTTCGGCTTCGGTATCAACGCTGAANAACACACAGAAATCTTCCTTGC	1680
1681	AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAGCAAGCGAAGGATTTAGCATTTCCC	1740
1681	AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAGCAAGCGAAGGATTTAGCATTTCCC	1740
1741	TGGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAAAACAGAAGGAATCTCACTTTGTGAG	1800
1741	TGGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAAAACAGAAGGAATCTCACTTTGTGAG	1800
1801	TGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCCTGAGAAAGAGTCTCTGAGAA	1860
1801	TGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCTCCTGAGAAAGAGTCTCTGAGAA	1860
1861	AGAGGATCAAGAGGAGGAAACCAAGGAGGGAAGGGTCCACTCTCTTTCAATTTTGAAGGC	1920
1861	AGAGGATCAAGAGGAGGAAACCAAGGAGGGAAGGGTCCACTCTCTTTCAATTTTGAAGGC	1920
1921	TTTTAACTGAGAATGGAGGCAACTTGTTATGTATTCGATAATAAGATCACGCTTTTGTACT	1980
1921	TTTTAACTGAGAATGGAGGCAACTTGTTATGTATTCGATAATAAGATCACGCTTTTGTACT	1980
1981	CTACTATCAAAAACTTATCAATAAATAAAAAAGTTTGGCGTTGTTTCTCC	2032
1981	CTACTATCAAAAACTTATCAATAAATAAAAAAGTTTGGCGTTGTTTCTCC	2032

RESULT 5	
AAAT76613	standard; cDNA to mRNA; 2032 BP.
ID	AAAT76613 standard; cDNA to mRNA; 2032 BP.
AC	AAAT76613;
XX	
XX	17-OCT-2003 (revised)
DT	29-DEC-1997 (first entry)
DT	
DE	
DE	Peanut allergen Ara hi cDNA clone P41b.
XX	
KW	Peanut; seed storage protein; allergen; a
KW	vaccine; anaphylactic shock; immunotherapy
KW	ELISA; analysis; Ara hi; ds.
XX	
OS	Arachis hypogaea; strain Florunner.
XX	
FH	Key Location/Qualifiers
FT	CDS 50..1930
FT	/*tag= a
FT	50..115
FT	/*tag= b
FT	116..1927
FT	/*tag= c
FT	polyA_signal 2005..2010
FT	/*tag= d
XX	
PN	WO9724139-A1.
XX	
PD	10-JUL-1997.
XX	
PF	23-SEP-1996; 96WO-US015222.
XX	
PR	29-DEC-1995; 95US-0009455P.
PR	04-MAR-1996; 96US-00610424.
XX	
PA	(UYAR-) UNIV ARKANSAS.
XX	
FI	Burks AW, Helm RM, Cockrell G, Stanley
XX	
XX	WPI; 1997-363453/33.
DR	P-PSDB; AAW22150.
DR	
XX	
PT	Peanut allergens Ara hi and Ara hII - use
PT	site monoclonal antibody based ELISA.
XX	
PS	Claim 31; Page 186-188; 354pp; English.
XX	
CC	This cDNA clone, designated P41b, codes for
CC	hi (AAW22150), which has multiple IGE bin
CC	It was amplified from peanut seed cDNA us
CC	based on an isolated Ara hi peptide (see
CC	significant homology with the vicilin fam
CC	other legumes. The gene is capable of pro
CC	prokaryotic cells that is recognised by s
CC	of individuals with peanut hypersensitivity
CC	AAW241164) can be used to raise monoclonal
CC	specific two-site Mab ELISA for the detec
CC	(claimed). IGE-binding Ara antigen epitop
CC	protect against allergic reactions to pea
CC	shock. (Updated on 17-OCT-2003 to standar
XX	
SQ	Sequence 2032 BP; 628 A; 473 C; 530 G; 40

[illegible]

QY	61	GGTTTCTC	CACTGATCGTGTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCA	120
Db				
QY	61	GGTTTCTC	CACTGATCGTGTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCA	120
Db				
QY	121	TGCCAAGT	CAATCACTTACAGAGAAAACAGAGAAACCCCTGCGCCACAGAGTGCTTCA	180
Db				
QY	121	TGCCAAGT	CATCACCTTACAGAGAAAACAGAGAAACCCCTGCGCCACAGAGTGCTTCA	180
Db				
QY	181	GAGTTGT	CAACAGGACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCGCTGCACCAA	240
Db				
QY	181	GAGTTGT	CAACAGGACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCGCTGCACCAA	240
Db				
QY	241	GCTCGAGT	ATGATCCTCTGTTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCA	300
Db				
QY	241	GCTCGAGT	ATGATCCTCTGTTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCA	300
Db				
QY	301	AGTTTCCC	CTCAGGGGACGGACACGTGGCGCCCAACCCCGAGACTACGATGATGACGG	360
Db				
QY	301	AGTTTCCC	CTCAGGGGACGGACACGTGGCGCCCAACCCCGAGACTACGATGATGACGG	360
Db				
QY	361	CGTCAACC	CCGAGAGAGGAGCGCGATGGGGACCAAGTGGAGCGACCTGGACCGAGGGGTGA	420
Db				
QY	361	CGTCAACC	CCGAGAGAGGAGCGCGATGGGGACCAAGTGGAGCGACCTGGACCGAGGGGTGA	420
Db				
QY	421	AAGAGAA	GAGACTGGAGACAACCAAGAGAAAGATTGGAGCGCACCAAGTCACTCAGCAGCC	480
Db				
QY	421	AAGAGAA	GAGACTGGAGACAACCAAGAGAAAGATTGGAGCGCACCAAGTCACTCAGCAGCC	480
Db				
QY	481	ACGAAAA	TAAGGCCCGAAGAGAGAGAGGAGAAACAAAGAGTGGGGAACACCAAGGTAGCCA	540
Db				
QY	481	ACGAAAA	TAAGGCCCGAAGAGAGAGAGGAGAAACAAAGAGTGGGGAACACCAAGGTAGCCA	540
Db				
QY	541	TGTGGG	GAAGAAACATCTCGGAACACCCCTTCTACTCCGCTCAAGCGGTTTAGCAC	600
Db				
QY	541	TGTGGG	GAAGAAACATCTCGGAACACCCCTTCTACTCCGCTCAAGCGGTTTAGCAC	600
Db				
QY	601	CCGCTAC	CGGGAACCAAAACGGTAGATCCGGGTCTCTGCAGAGGTTTGACCAAGGTCAAG	660
Db				
QY	601	CCGCTAC	CGGGAACCAAAACGGTAGATCCGGGTCTCTGCAGAGGTTTGACCAAGGTCAAG	660
Db				
QY	661	GCAGTTT	CAGATCTCCAGAAATCACCGTTATGTGCGAGTTCGAGGCGCAACCTAACTCT	720
Db				
QY	661	GCAGTTT	CAGAAATCTCCAGAAATCACCGTTATGTGCGAGTTCGAGGCGCAACCTAACTCT	720
Db				
QY	721	TGTTCTT	CCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGAGGSCAAGCCAC	780
Db				
QY	721	TGTTCTT	CCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGAGGSCAAGCCAC	780
Db				
QY	781	CGTGAC	CGGTAGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACT	840
Db				
QY	781	CGTGAC	CGGTAGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACT	840
Db				
QY	841	CAGAAAT	CCCATCCGGTTTCATTTCTTAATCTTTGAAACCGCCATGACCAACCACTCAG	900
Db				
QY	841	CAGAAAT	CCCATCCGGTTTCATTTCTTAATCTTTGAAACCGCCATGACCAACCACTCAG	900
Db				
QY	901	AGTAGT	AAAAATCTCCATGCCCGTTAACACACCCCGGCTTGGAGTTTCTCCCGGC	960
Db				
QY	901	AGTAGT	AAAAATCTCCATGCCCGGTACAAACCCCGGCGAGTTTGGAGTTTCTCCCGGC	960
Db				
QY	961	GAGCAG	CCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGGCGC	1020
Db				
QY	961	GAGCAG	CCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGGCGC	1020
Db				
QY	1021	CTTCAAT	CGGAAATTCATAGATACGGAGGTGCTGTTAGAGAAATGCAAGAGGTGA	1080
Db				
QY	1021	CTTCAAT	CGGAAATTCATAGATACGGAGGTGCTGTTAGAGAAATGCAAGAGGTGA	1080
Db				
QY	1081	GCAAGG	AGAGAGGCGAGGCGATGGAGTACTCGGAGTACTGAGAACTAAGAGGAGT	1140
Db				
QY	1081	GCAAGG	AGAGAGGCGAGGCGATGGAGTACTCGGAGTACTGAGAACTAAGAGGAGT	1140
Db				

QY	1141	GATAGT	CAAGGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCACGCTAAATCGGTCTC	1200
Db				
QY	1141	GATAGT	CAAGGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCACGCTAAATCGGTCTC	1200
Db				
QY	1201	AAAGAA	AGGCTCCGAAAGAGGAGGAGATATCACCAACCAATCAACTTTGAGAAAGGGA	1260
Db				
QY	1201	AAAGAA	AGGCTCCGAAAGAGGAGGAGATATCACCAACCAATCAACTTTGAGAAAGGGA	1260
Db				
QY	1261	GCCGAT	CTTTCTAACTTTGGGAGCTTATTTGAGGTGAAGCCAGACAAGAGAACCC	1320
Db				
QY	1261	GCCGAT	CTTTCTAACTTTGGGAGCTTATTTGAGGTGAAGCCAGACAAGAGAACCC	1320
Db				
QY	1321	CCAGTTC	CAGGACCTGGACATGCTCACCTGTGTAGAGATCAAAGAGGAGCTTTGAT	1380
Db				
QY	1321	CCAGTTC	CAGGACCTGGACATGCTCACCTGTGTAGAGATCAAAGAGGAGCTTTGAT	1380
Db				
QY	1381	GCTCCCA	CACTTCAACTCAAAGGCCATGGTTATCGTCTGTCGAACAAAGGACCTGGAAA	1440
Db				
QY	1381	GCTCCCA	CACTTCAACTCAAAGGCCATGGTTATCGTCTGTCGAACAAAGGACCTGGAAA	1440
Db				
QY	1441	CCTTGA	ACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGGAAGAGAGA	1500
Db				
QY	1441	CCTTGA	ACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGGAAGAGAGA	1500
Db				
QY	1501	GGACGA	ACGAAGAGGAGGGAAGTAACAGAGAGGTGGTGTAGGTACACAGCGAGGTT	1560
Db				
QY	1501	GGACGA	ACGAAGAGGAGGGAAGTAACAGAGAGGTGGTGTAGGTACACAGCGAGGTT	1560
Db				
QY	1561	GAAGAA	GGCGATGTTTCATGCGCAGAGCTCATCCAGTAGCCATCAACGCTTCTC	1620
Db				
QY	1561	GAAGAA	GGCGATGTTTCATGCGCAGAGCTCATCCAGTAGCCATCAACGCTTCTC	1620
Db				
QY	1621	CGAATC	CCATCTGCTTGGCTTCAACCGCTGAAACCAACACAGAACTTCTCTTGC	1680
Db				
QY	1621	CGAATC	CCATCTGCTTGGCTTCAACCGCTGAAACCAACACAGAACTTCTCTTGC	1680
Db				
QY	1681	AGTGAT	AAGGACAAATGTGTAGACCAAGAGCAAGCAAGGAGTTTAGCATTTCC	1740
Db				
QY	1681	AGTGAT	AAGGACAAATGTGTAGACCAAGAGCAAGCAAGGAGTTTAGCATTTCC	1740
Db				
QY	1741	TGGTTC	GGGTGAAACGAGTTGAGAGCTCATCAAAACCAAGAGGATCTCACTTTGTGAG	1800
Db				
QY	1741	TGGTTC	GGGTGAAACGAGTTGAGAGCTCATCAAAACCAAGAGGATCTCACTTTGTGAG	1800
Db				
QY	1801	TGCTCG	TCTCAATCTCAATCTCCGCTCGTCTCTGAGAAAGAGTCTCTTGAGAA	1860
Db				
QY	1801	TGCTCG	TCTCAATCTCAATCTCCGCTCGTCTCTGAGAAAGAGTCTCTTGAGAA	1860
Db				
QY	1861	AGAGAT	CAAGGAGGAAAAACCAAGGAGGAGGTCCTCTTTCAATTTTGAAGGC	1920
Db				
QY	1861	AGAGAT	CAAGGAGGAAAAACCAAGGAGGAGGTCCTCTTTCAATTTTGAAGGC	1920
Db				
QY	1921	TTTTAA	CTGAGATGAGGAGCACTTGTATGATCGATAAAGATCAAGCTTTGTACT	1980
Db				
QY	1921	TTTTAA	CTGAGATGAGGAGCACTTGTATGATCGATAAAGATCAAGCTTTGTACT	1980
Db				
QY	1981	CTACTAT	CCAAAACTTATCAATAAATAAAAAAGTTTGTGCTGTTTCTCC	2032
Db				
QY	1981	CTACTAT	CCAAAACTTATCAATAAATAAAAAAGTTTGTGCTGTTTCTCC	2032
Db				

RESULT 6

AAZ06382

ID AAZ06382 standard; DNA; 1930 BP.

XX AAZ06382;

XX 17-OCT-2003 (revised)

XX 09-NOV-1999 (first entry)

XX Peanut allergen, Ara h 1.

DE Peanut allergen, Ara h 1.

XX allergy; immune response; transgenic; allergen; epitope;

immunoglobulin E; Ig E; binding site; peanut; ds.

Arachis hypogaea.

Key Location/Qualifiers
CDS 50..1930
/*tag= a
/product= "Ara h 1"

W09938978-A1.

05-AUG-1999.

29-JAN-1999; 99WO-US002031.

31-JAN-1998; 98US-0073283P.

13-FEB-1998; 98US-0074590P.

13-FEB-1998; 98US-0074624P.

13-FEB-1998; 98US-0074633P.

27-AUG-1998; 98US-00141220.

(UYAR-) UNIV ARKANSAS.

(UYNV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

(SOSI/) SOSIN H.

Sosin H, Bannon GA, Burks AW, Sampson HA;

WPI; 1999-479189/40.

P-PSDB; AAY15244.

Modified allergen with reduced IgE binding, useful for treating e.g. allergies.

Disclosure; Page 32-33; 46pp; English.

This is the nucleotide sequence of the Ara h 1 protein from Arachis hypogaea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding epitopes, four of which are immunodominant (AAY15247, AAY15249, AAY15250 and AAY15263). By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response. (Updated on 17 -OCT-2003 to standardise OS field)

Sequence 1930 BP; 595 A; 455 C; 514 G; 366 T; 0 U; 0 Other;

Query Match 95.0%; Score 1930; DB 2; Length 1930;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATAATCATATATATTCATCAATCATCTATATTAAGTAGTAGCAGGAGCAATGAGGGAG 60

1 AATAATCATATATATTCATCAATCATCTATATTAAGTAGTAGCAGGAGCAATGAGGGAG 60

61 GGTTCCTCCAGTATGCTGTGTAGGATCCTGTGCTGGCTTCAGTTCTGCAACGCA 120

61 GGTTCCTCCAGTATGCTGTGTAGGATCCTGTGCTGGCTTCAGTTCTGCAACGCA 120

121 TGCCAAGTATCATACCTTACCAGAAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA 180

121 TGCCAAGTATCATACCTTACCAGAAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA 180

181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCAATGCGATGCTGCAACCA 240

181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCAATGCGATGCTGCAACCA 240

241 GCTCGAGTATGATCCTGCTGTCTATGATCTCTCGAGGACACATGCGGACCAACCA 300

241 GCTCGAGTATGATCCTGCTGTCTATGATCTCTCGAGGACACATGCGGACCAACCA 300

301 ACCTTCCCTCCAGGGAGCGGACAGTGGCCGCCCAACCCGGAGACTACGATGATGACCG 360

301 ACCTTCCCTCCAGGGAGCGGACAGTGGCCGCCCAACCCGGAGACTACGATGATGACCG 360

301 ACCTTCCCTCCAGGGAGCGGACAGTGGCCGCCCAACCCGGAGACTACGATGATGACCG 360

QY	361	CCGTCAACCCCGAAGAGAGAGAGGCCGATGGGGACACAGCTGGACCCAGGGAGCGTGA	420
DB	361	CCGTCAACCCCGAAGAGAGAGAGGCCGATGGGGACACAGCTGGACCCAGGGAGCGTGA	420
QY	421	AAGAGAAGAAGACTGGAGACAAACCAAGAGAGATTGGAGCGCACCAAGTCAATCAGCAGCC	480
DB	421	AAGAGAAGAAGACTGGAGACAAACCAAGAGAGATTGGAGCGCACCAAGTCAATCAGCAGCC	480
QY	481	ACGGAAATTAAGGCCCGAAG	540
DB	481	ACGGAAATTAAGGCCCGAAG	540
QY	541	TGTGAGGGAAGAAAATCTCGGAAACAAACCTTTCTAATCTTCCCTCAAGGGGTTTACAC	600
DB	541	TGTGAGGGAAGAAAATCTCGGAAACAAACCTTTCTAATCTTCCCTCAAGGGGTTTACAC	600
QY	601	CCGTACGGAACCAAAACCGTAGGATCGGGTCTCGAGAGGTTTGAACCAAGGTCAAG	660
DB	601	CCGTACGGAACCAAAACCGTAGGATCGGGTCTCGAGAGGTTTGAACCAAGGTCAAG	660
QY	661	GCAGTTTCAGAAATCTCCAGATCACCGTATTTGTCAGATCGAGGCCCAACCTAACACTCT	720
DB	661	GCAGTTTCAGAAATCTCCAGATCACCGTATTTGTCAGATCGAGGCCCAACCTAACACTCT	720
QY	721	TGTTCTTCCAGACGCTGATGATTAACATCTTGTATTCAGCAAGGGCAAGCCAC	780
DB	721	TGTTCTTCCAGACGCTGATGATTAACATCTTGTATTCAGCAAGGGCAAGCCAC	780
QY	781	CGTGACCGTAGCAATTAACAGAAAGAGCTTTAATCTTTCAGAGGGCCATCACT	840
DB	781	CGTGACCGTAGCAATTAACAGAAAGAGCTTTAATCTTTCAGAGGGCCATCACT	840
QY	841	CAGAAATCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACCAACAGAACCTCAG	900
DB	841	CAGAAATCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACCAACAGAACCTCAG	900
QY	901	AGTAGTAAATCTCCATGCGCTTAAACACACCCCGCCAGTTTGAAGATTTCTTCCGGC	960
DB	901	AGTAGTAAATCTCCATGCGCTTAAACACACCCCGCCAGTTTGAAGATTTCTTCCGGC	960
QY	961	GAGCAGCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGCCG	1020
DB	961	GAGCAGCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGCCG	1020
QY	1021	CTTCAATCGGAATTCATGAGATACGAGGGTGTCTTTAGAGAGAAATGCAAGGAGTGA	1080
DB	1021	CTTCAATCGGAATTCATGAGATACGAGGGTGTCTTTAGAGAGAAATGCAAGGAGTGA	1080
QY	1081	GCAAGAGAGAGAGGGCAGAGCGATGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT	1140
DB	1081	GCAAGAGAGAGAGGGCAGAGCGATGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT	1140
QY	1141	GATAGTCAAGGTGTCAGAGGACAGCTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTC	1200
DB	1141	GATAGTCAAGGTGTCAGAGGACAGCTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTC	1200
QY	1201	AAAGAAAGGCTCCGAGAGAGAGAGATATCACCAACCAATCACTTTCAGAGAGAGCGA	1260
DB	1201	AAAGAAAGGCTCCGAGAGAGAGAGATATCACCAACCAATCACTTTCAGAGAGAGCGA	1260
QY	1261	GCCGATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACCAAGAAACCC	1320
DB	1261	GCCGATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACCAAGAAACCC	1320
QY	1321	CCAGCTTCAGGACCTGGAATCTACCTGTGTAGAGATCAAGAGAGGAGCTTTGAT	1380
DB	1321	CCAGCTTCAGGACCTGGAATCTACCTGTGTAGAGATCAAGAGAGGAGCTTTGAT	1380
QY	1381	GCTCCCACTTCACTCAAGGCCATGTTATCGTCGTCAACCAAGGAACTGGA	1440
DB	1381	GCTCCCACTTCACTCAAGGCCATGTTATCGTCGTCAACCAAGGAACTGGA	1440

QY	1441	CCTTGAACCTGGTGGCTGTGAAGAAAAGAGCAACAAACAGAGGGGACGGCGGGAAGAAGAGGA	1500
Db	1441	CCTTGAACCTGGTGGCTGTGAAGAAAAGAGCAACAAACAGAGGGGACGGCGGGAAGAAGAGGA	1500
QY	1501	GGACGAAGACGAAGAAGAGAGGGAAGTAACACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
Db	1501	GGACGAAGACGAAGAAGAGAGGGAAGTAACACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
QY	1561	GAAGGAAGCGGATGTGTTTCATCATGCGACGAGCTCATCCAGTAGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Db	1561	GAAGGAAGCGGATGTGTTTCATCATGCGACGAGCTCATCCAGTAGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
QY	1621	CGAACTCCATCTGCTTTGGCTTCGGTATCAACGCTGAAAAACACCAAGCAATCTTCCTGC	1680
Db	1621	CGAACTCCATCTGCTTTGGCTTCGGTATCAACGCTGAAAAACACCAAGCAATCTTCCTGC	1680
QY	1681	AGTGATAGGACAATGTGATAGACACAGATAGAGAAAGCAAGCGAAGGATTTAGCATTC	1740
Db	1681	AGTGATAGGACAATGTGATAGACACAGATAGAGAAAGCAAGGATTTAGCATTC	1740
QY	1741	TGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGGAAGTAATCTCACTTTGTGAG	1800
Db	1741	TGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGGAAGTAATCTCACTTTGTGAG	1800
QY	1801	TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGTGAAGAGAGTCTCCTGAGAA	1860
Db	1801	TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGTGAAGAGAGTCTCCTGAGAA	1860
QY	1861	AGAGGATCAAGAGGAGGAACCAAGGAGGAGGTCACCTCCTTTGAAGGC	1920
Db	1861	AGAGGATCAAGAGGAGGAACCAAGGAGGAGGTCACCTCCTTTGAAGGC	1920
QY	1921	TTTTAACTGA 1930	
Db	1921	TTTTAACTGA 1930	
RESULT 7			
AAT76612			
ID	AAT76612 standard; cDNA to mRNA; 1949 BP.		
XX	AAT76612;		
AC			
XX			
DT	17-OCT-2003 (revised)		
DT	29-DEC-1997 (first entry)		
XX	Peanut allergen Ara hI cDNA clone P17.		
KW	Peanut; seed storage protein; allergen; allergy; hypersensitivity;		
KW	vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;		
KW	ELISA; analysis; Ara hI; ds.		
OS	Arachis hypogaea; strain Flurunner.		
XX			
FH	Key		
FT	CDS	Location/Qualifiers	
FT	3..1847	/*tag= a	
FT	3..68	/*tag= b	
FT	65..1844	/*tag= c	
FT	1918..1923	/*tag= d	
FT			
XX	WO9724139-A1.		
XX	10-JUL-1997.		
XX	23-SEP-1996; 96WO-US015222.		
XX	29-DEC-1995; 95US-0009455P.		
XX	04-MAR-1996; 96US-0061042A.		

PA	(UYAR-) UNIV ARKANSAS.	
XX		
PI	Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;	
XX		
DR	WPI; 1997-363453/33.	
DR	P-PSDB; AAW22149.	
XX		
PT	Peanut allergens Ara hI and Ara hII - used for vaccination and in two-	
XX	site monoclonal antibody based ELISA.	
PS	Claim 31; Page 183-185; 354pp; English.	
XX		
CC	This cDNA clone, designated p17, codes for the major peanut allergen Ara	
CC	hI (AAW22149), which has multiple IGE binding epitopes (see AAW24165-87).	
CC	It was amplified from peanut seed cDNA using a primer (see AAT76616)	
CC	based on an isolated Ara hI peptide (see AAW24206). The sequence shows	
CC	significant homology with the vicilin family of seed storage proteins of	
CC	other legumes. The gene is capable of producing a protein product in	
CC	prokaryotic cells that is recognised by serum IGE from a large proportion	
CC	of individuals with peanut hypersensitivity. Ara hI and Ara hII (see	
CC	AAW24164) can be used to raise monoclonal antibodies which are used in a	
CC	specific two-site Mab ELISA for the detection of Ara hI or Ara hII	
CC	(claimed). IGE-binding Ara antigen epitopes may be used in vaccines to	
CC	protect against allergic reactions to peanut allergens, e.g. anaphylactic	
CC	shock. (Updated on 17-OCT-2003 to standardise OS field)	
XX		
SQ	Sequence 1949 BP; 599 A; 455 C; 517 G; 378 T; 0 U; 0 Other;	
Query Match 86.2%; Score 1752.4; DB 2; Length 1949;		
Best Local Similarity 95.0%; Pred. No. 0;		
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;		
QY	48	CAATGAGAGGAGGGTTCCTCCACTGATGCTGTGTCTAGGATCCTGTCTCGGCTTCAG 107
Db	1	CAATGAGAGGAGGGTTCCTCCACTGATGCTGTGTCTAGGATCCTGTCTCGGCTTCAG 60
QY	108	TTTCTGCAACGATGCCAAGTCATCCTTACGAGAAAGAAACAGAGAACCCCTCGGCC 167
Db	61	TTTCTGCAACGAGGCCAAGTCACC-----TTACCGAAAGAACAGAGAACCCCTCGGCC 114
QY	168	AGAGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGT 227
Db	115	AGAGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGT 174
QY	228	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTG 287
Db	175	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATG-----ACACTG 222
QY	288	GCACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCGCCAAACCCGAGACT 347
Db	223	GGCCACCAACCAACGTCACCTCCAGGGAGCGGACACGTGGCGCCAAACCCGAGACT 282
QY	348	ACGATGATGACCGCGTCAACCCCGAAGAGAGGAGGCGGATGGGACCGAGTGGAC 407
Db	283	ACGATGATGACCGCGTCAACCCCGAAGAGAGGAGGCGGATGGGACCGAGTGGAC 342
QY	408	CGAGGAGCGTGAAAGAGAAGACTGGAGACCAACCAAGAGAAGATTGGAGGCGACCAA 467
Db	343	CGAGGAGCGTGAAAGAGAAGACTGGAGACCAACCAAGAGAAGATTGGAGGCGACCAA 402
QY	468	GTATCAGACGCCACGGAAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGATGGGAA 527
Db	403	GTATCAGACGCCACGGAAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGATGGGAA 462
QY	528	CACAGGTAGCCATGTGAGGGAAGAAACATCTCCGAACAAACCTTTCTACTTCCCGTCAA 587
Db	463	CACAGGTAGCCATGTGAGGGAAGAAACATCTCCGAACAAACCTTTCTACTTCCCGTCAA 522
QY	588	GGCGGTTTAGCACCCGCTACCGGAACCAAAACCGGTAGGATCCGGGTCTTCAGAGGTTTG 647
Db	523	GGCGGTTTAGCACCCGCTACCGGAACCAAAACCGGTAGGATCCGGGTCTTCAGAGGTTTG 582
QY	648	ACCAAAGGTCAAGGCGAGTTTCAGAAATCTCCAGAAATCACCGGTATTGTGCAGATCGAGGCCA 707

||||| 583 ACCAAGGTCAGAGCAGTTTTCAGAAATCTCCAGAAATCACCGTATTTGTGCAGATCAGGCCA 642
Qy 708 AACCTAACACTCTTTGTTCTTCCCAAGCAGCGTGATGCTGATAAATCTCTTTGTTATCCAGC 767
Db 643 GACCTAACACTCTTTGTTCTTCCCAAGCAGCGTGATGCTGATAAATCTCTTTGTTATCCAGC 702
Qy 768 AAGGGCAAGCCACCGTCAGCGTAGCAATGGCAATACAGAAAGAGGTTTAAATCTTTGAGC 827
Db 703 AAGGCAAGCCACCGTGACCGTAGCAATGGCAATACAGAAAGAGGTTTAAATCTTTGAGC 762
Qy 828 AAGGCCATGCACTCAGAAATCCCATCCCGTTTCAATTTCTTACATCTTTGAAACCGCCATGACA 887
Db 763 AAGGCCATGCACTCAGAAATCCCATCCCGTTTCAATTTCTTACATCTTTGAAATCGACATGACA 822
Qy 888 ACAGAAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACACCGCGCAGTTTGGAG 947
Db 823 ACCAGAACTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACACCGCGCGCAGTTTGGAG 882
Qy 948 ATTTCTTCCCGCGGAGCAGCGAGACCAATCATCTTACTTGCAGGGCTTCCAGCAGGAATA 1007
Db 883 ATTTCTTCCCGCGGAGCAGCGAGACCAATCATCTTACTTGCAGGGGATTCAGCAGGAATA 942
Qy 1008 CGTTGAGGCGCGCTTCAATCGCGAATTCATGAGATACGAGGGTCTGTAGAAAGAGA 1067
Db 943 CTTTGGAGCGCGCTTCAATCGCGAATTCATGAGATACGAGGGTCTGTAGAAAGAGA 1002
Qy 1068 ATGAGGAGGTGAGCAAGAGAGAGAGGGGAGAGGGGATGAGTACTCGGAGTAGTGAGA 1127
Db 1003 ATGAGGAGGAGAGCAAGAGAGAGAGGGGAGAGGGGATGAGTACTCGGAGTAGTG-- 1059
Qy 1128 ACAATGAGGAGTAGTAGTCAAAAGTGTCAAGAGGACGTTGAAGAACTTACTTAAGCAG 1187
Db 1060 ATAATGAGGAGTAGTAGTCAAAAGTGTCAAGAGGACGTTCAAGAACTTACTTAAGCAG 1119
Qy 1188 CTAATCCGTCCTCAAGAAAGGCTCCGAGAGAGGAGATATCACCAACCCCAATCAACT 1247
Db 1120 CTAATCCGTCCTCAAGAAAGGCTCC--GAAGAGGAGATATCACCAACCCCAATCAACT 1176
Qy 1248 TGAGAGAGCGGAGCGCGATCTTTTCTAACTTTTGGAAAGTTATTGAGGTGAAGCCAG 1307
Db 1177 TGAGAGAGCGGAGCGCGATCTTTTCTAACTTTTGGAGGTTATTGAGGTGAAGCCAG 1236
Qy 1308 ACAAGAGAAACCCGAGCTTCAGACCTGACATGATGCTCAGCTGTGTAGAGATCAAG 1367
Db 1237 ACAAGAGAAACCCGAGCTTCAGACCTGACATGATGCTCAGCTGTGTAGAGATCAAG 1296
Qy 1368 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATGCTCGTCAACA 1427
Db 1297 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATGCTCGTCAACA 1356
Qy 1428 AAGGAACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGC 1487
Db 1357 AAGGAACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGC 1416
Qy 1488 GGGAA-----GAAGAGGAGCAAGACAGCAAGAGAGGAGGAGTAACAGAGAG 1538
Db 1417 GGGAAACAGAGTGGGAGAGAGAGGAGGAGATGAAGAGAGGAGGAGGAGTAACAGAGAG 1476
Qy 1539 TGGTAGGTACAGAGGAGGTTGAAGGAGGCGATGTTTCAATGCGCAGCAGCTCATC 1598
Db 1477 TGGTAGGTACAGAGGAGGTTGAAGGAGGCGATGTTTCAATGCGCAGCAGCTCATC 1536
Qy 1599 CAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1596
Qy 1659 ACAACACAGAAATCTTCTTGCAGGTGATGAAGCAATGTGATAGACAGATAGAAGC 1718
Db 1597 ACAACACAGAAATCTTCTTGCAGGTGATGAAGCAATGTGATAGACAGATAGAAGC 1656
Qy 1719 AAGGAGGAGTTAGCATTCCTCGGCTGAGTCAAGTTGAGAGGCTCATCAAAACC 1778
|||||

Db 1657 AAGCGAAGGATTTAGCATTCCTGGTTCCGGTGAAACAAGTTGAGAAGCTCATCAAAACC 1716
Qy 1779 AGAAGAAATCTCACTTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCGTCTC 1838
Db 1717 AGAGGAGGATCTCACTTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTC----- 1765
Qy 1839 CTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGGTC 1898
Db 1766 -----GTCCTCTGAAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGGTC 1815
Qy 1899 CACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAATGAGGCAACTTGTATTATCTATCGAT 1958
Db 1816 CACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAATGAGGAAACTTGTATTATCTATCAT 1875
Qy 1959 AATAAGATCAGCGTTTGTACTACTACTATCCAAAAAATCTATCAATAAATAAAACGTTTG 2018
Db 1876 AATAAGATCAGCGTTTGTAAATCTACTATCCAAAAAATCTATCAATAAATAAAACGTTTG 1935
Qy 2019 TCGTTTGTCTTCC 2032
Db 1936 TCGTTTGTCTTCC 1949
RESULT 8
ABX70604
ID ABX70604 standard; cDNA; 1949 BP.
XX
AC ABX70604;
AC AC
XX XX
DT 26-MAR-2003 (first entry)
XX
DE Peanut Ara h1 cDNA clone P17.
XX
KW Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; IgE binding site;
KW anaphylactic food allergen; anti-allergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabin PA, Shin DS, Stanley JS;
XX
DR WPI: 2003-018765/01.
DR P-PSDB; ABU52413.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Example 3; Fig 11; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified

CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a cDNA
CC encoding a peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 1949 BP; 599 A; 455 C; 517 G; 378 T; 0 U; 0 Other;

Query Match 86.2%; Score 1752.4; DB 8; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;

QY	48	CAATGAGAGGGGTTCTCCACTGATGCTGTGTAGGGATCCCTTGTCTCGGCTTCAG	107
DB	1	CAATGAGAGGGGTTCTCCACTGATGCTGTGTGTGGGATCCTTGTCTCGGCTTCAG	60
QY	108	TTTCTGCAACGATGCAAGTCAATCACTTACAGAGAAACACAGAGAACCCCTGGGCC	167
DB	61	TTTCTGCAACGAGGCCAAGTCACC-----TTACCGGAACACAGAGAACCCCTGGGCC	114
QY	168	AGAGGTGCTCCAGAGTTGTCACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGT	227
DB	115	AGAGGTGCTCCAGAGTTGTCACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGT	174
QY	228	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGACACACTG	287
DB	175	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTGTCTATG-----ACACTG	222
QY	288	GCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCGAGACT	347
DB	223	GGGCCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCGAGACT	282
QY	348	ACGATGATGACCGCTGACCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	407
DB	283	ACGATGATGACCGCTGACCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	342
QY	408	CGAGGAGCGCTGAAAGAGAAAGAGCTGGAGACCAACCAAGAGAAAGTTGGAGCGACCA	467
DB	343	CGAGGAGCGCTGAAAGAGAAAGAGCTGGAGACCAACCAAGAGAAAGTTGGAGCGACCA	402
QY	468	GTCAATCAGCAGCCACGGAATAAAGCCCGAAGGAAGAGAGAGAGAGAGAGAGAGAG	527
DB	403	GTCAATCAGCAGCCACGGAATAAAGCCCGAAGGAAGAGAGAGAGAGAGAGAGAGAG	462
QY	528	CACCGTAGCATGTGAGGGAGAAACATCTCGGAACACCCCTTCTACTTCCCGTCAA	587
DB	463	CACCGTAGCATGTGAGGGAGAAACATCTCGGAACACCCCTTCTACTTCCCGTCAA	522
QY	588	GGCGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTTCGACAGAGTTG	647
DB	523	GGCGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTTCGACAGAGTTG	582
QY	648	ACCAAGGTCAAGCGAGTTTCAAGATCTCCAGATACCGTATGTGACAGATCGAGGCCA	707
DB	583	ACCAAGGTCAAGCGAGTTTCAAGATCTCCAGATACCGTATGTGACAGATCGAGGCCA	642
QY	708	AACCTAACAACCTGTGTTCTTCCAGCAGCGCTGATGAACATCTCTGTTTATCCAGC	767
DB	643	GACCTAACAACCTGTGTTCTTCCAGCAGCGCTGATGAACATCTCTGTTTATCCAGC	702
QY	768	AAGGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACG	827
DB	703	AAGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACG	762
QY	828	AGGGCATTGCACTCAGAAATCCATCGGTTTCATTTCTTACATCTTGAAACCCGCCATGACA	887
DB	763	AGGGCATTGCACTCAGAAATCCATCGGTTTCATTTCTTACATCTTGAAATCGACATGACA	822
QY	888	ACCAGAACCTCAGAGTAGCTAAATCTCCATCCCGGTTAAACACACCCGCCAGTTTGAGG	947
DB	823	ACCAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAGG	882

QY	948	ATTTCTTCCCGCGAGCGAGCCGAGACCAATCATCTCTACTTGCAGGCTTTCAGCAGGAATA	1007
DB	893	ATTTCTTCCCGCGAGCGAGCCGAGACCAATCATCTCTACTTGCAGGATTCAGCAGGAATA	942
QY	1008	CGTTGGAGCGCCCTTCAATGCGGAATTCATGAGATACGGAGGTTGCTTTAGAGAGA	1067
DB	943	CGTTGGAGCGCCCTTCAATGCGGAATTCATGAGATACGGAGGTTGCTTTAGAGAGA	1002
QY	1068	ATGCAAGGAGTTCAGCAAGGAGGAGCGAGCGGATGGAGTACTCGGAGTACTGAGA	1127
DB	1003	ATGCAAGGAGGAGCAAGGAGGAGCGGAGCGGATGGAGTACTCGGAGTACTGAGA	1059
QY	1128	ACAATCAAGGAGTTCATGAGTCAAAAGTGTCAAAAGGACACGTTGAAGAACTTACTAAGCAG	1187
DB	1060	ATAATGAAGGAGTTCATGAGTCAAAAGTGTCAAAAGGACACGTTCAAGAACTTACTAAGCAG	1119
QY	1188	CTAAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGAGAGATATCAACCAACCAATCAACT	1247
DB	1120	CTAAATCCGTCTCAAGAAAGGCTCC--GAAGAGAAAGATATCAACCAACCAATCAACT	1176
QY	1248	TGAGAGAGGCGAGCCCGATCTTTCTAAACCTTTGGGAAGTTATTTCGAGGTGAAGCCAG	1307
DB	1177	TGAGAGAGTGGCGAGCCCGATCTTTCTAAACCTTTGGGAAGTTATTTCGAGGTGAAGCCAG	1236
QY	1308	ACAAGAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG	1367
DB	1237	ACAAGAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG	1296
QY	1368	AAGGAGCTTTGATGCTCCACACCTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACA	1427
DB	1297	AAGGAGCTTTGATGCTCCACACCTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACA	1356
QY	1428	AAGGAACCTGGAACCTTCGAACCTCGTGGCTGTGAAGAAAGAGCAACACAGAGGGGACCGC	1487
DB	1357	AAGGAACCTGGAACCTTCGAACCTCGTGGCTGTGAAGAAAGAGCAACACAGAGGGGACCGC	1416
QY	1488	GGGA-----GAAGAGGAGGACGAAGACGAAGAGAGAGAGGAGGAGTAACAGAGAGG	1538
DB	1417	GGGAACGAAGTGGGAAGAGAGGAGGAAGATGAAGAGAGAGGAGGAGTAACAGAGAGG	1476
QY	1539	TGCGTAGGTACACAGCAGGTTGAAGGAAGGGATGTGTTTCATCATGCGCAGCAGCTCATC	1598
DB	1477	TGCGTAGGTACACAGCAGGTTGAAGGAAGGGATGTGTTTCATCATGCGCAGCAGCTCATC	1536
QY	1599	CAGTAGCCATCAACGCTTCTCCGAACCTCTGCTTGGCTTCGGTATCAACGCTGAAA	1658
DB	1537	CAGTAGCCATCAACGCTTCTCCGAACCTCTGCTTGGCTTCGGTATCAACGCTGAAA	1596
QY	1659	ACAACACAGAACTCTTTCAGGTGATAAGGACAAATGTGATAGACAGATAGAGAAGC	1718
DB	1597	ACAACACAGAACTCTTTCAGGTGATAAGGACAAATGTGATAGACAGATAGAGAAGC	1656
QY	1719	AAGCGAAGATTTAGCATTCCTGGGTGGGTGAACAAGTTGAGAGCTCATCAAAAACC	1778
DB	1657	AAGCGAAGATTTAGCATTCCTGGGTGGGTGAACAAGTTGAGAGCTCATCAAAAACC	1716
QY	1779	AGAAGGAATCTCACCTTTCGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTC	1838
DB	1717	AGAAGGAATCTCACCTTTCGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTC	1765
QY	1839	CTGAGAAAGAGTCTCTCGAGAAAGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGTC	1898
DB	1766	-----GTCTCTGAAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGTC	1815
QY	1899	CATCTCTTCAATTTTGAAGCTTTTAACTGAGAAATGAGGCACTTGTATGATCAT	1958
DB	1816	CATCTCTTCAATTTTGAAGCTTTTAACTGAGAAATGAGGCACTTGTATGATCAT	1875
QY	1959	AATAAGATCAGCTTTTGTACTATCCAAAACCTTATCAATAAATAAAGGTTTG	2018
DB	1876	AATAAGATCAGCTTTTGTATCTACTATCCAAAACCTTATCAATAAATAAAGGTTTG	1935
QY	2019	TGCGTGTGTTCTCC	2032

Db 1936 TGCCTGTTCTTC 1949

|||||

RESULT 9

ADG27463

ID ADG27463 standard; cDNA; 1949 BP.

XX AC ADG27463;

XX DT 26-FEB-2004 (first entry)

XX DE Peanut Ara h1 cDNA clone P17.

XX KW Peanut; plant; ss; gene; allergen; Ara h1; Ara h2; Ara h3;

XX KW glycinin A2B1a; Jug n1; antiallergic; vulnery;

XX KW anaphylactic food allergen; IgE; allergy; wound.

XX OS Arachis hypogaea.

XX PN US2003202980-A1.

XX PD 30-OCT-2003.

XX PF 18-MAR-2002; 2002US-00100303.

XX PR 29-DEC-1995; 95US-0009455P.

XX PR 23-SRP-1996; 96US-00717933.

XX PR 31-JAN-1998; 98US-0073283P.

XX PR 13-FEB-1998; 98US-0074590P.

XX PR 13-FEB-1998; 98US-0074624P.

XX PR 13-FEB-1998; 98US-0074633P.

XX PR 27-AUG-1998; 98US-00106872.

XX PR 27-AUG-1998; 98US-00141220.

XX PR 13-NOV-1998; 98US-00191593.

XX PR 23-JAN-1999; 99US-00240557.

XX PR 23-JAN-1999; 99US-00241101.

XX PR 11-FEB-1999; 99US-00248673.

XX PR 11-FEB-1999; 99US-00248674.

XX PR 02-MAR-1999; 99US-0122450P.

XX PR 02-MAR-1999; 99US-0122452P.

XX PR 02-MAR-1999; 99US-0122560P.

XX PR 02-MAR-1999; 99US-0122566P.

XX PR 11-MAR-1999; 99US-00267719.

XX PR 28-JAN-2000; 2000US-00494096.

XX PR 16-MAR-2001; 2001US-0276822P.

XX (CAPL/) CAPLAN M J.

XX (SOSI/) SOSIN H B.

XX (SAMP/) SAMPSON H.

XX (BANN/) BANNON G A.

XX (BURK/) BURKS A W.

XX (COCK/) COCKRELL G.

XX (COMP/) COMPADRE C M.

XX (CONN/) CONNAUGHTON C.

XX (KING/) KING N E.

XX (KOPP/) KOPPER R A.

XX (MALE/) MALEKI S J.

XX (RABJ/) RABJOHN P A.

XX (SHIN/) SHIN D S.

XX (STAN/) STANLEY J S.

XX Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;

XX Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

XX Rabjohn PA, Shin DS, Stanley JS;

XX WPI: 2003-875632/81.

XX P-PSDB; ADG27465.

XX New modified anaphylactic food allergen comprising a cysteine residue

XX PT which has been modified so that it cannot participate in the disulfide

Db 523 |||||GGCGGTTTAGCACCCGCTAGCGGAACCAAAACGGTAGGATCCGCGCTCTCGACAGGTTTG 582
Qy 648 ACCAAAGGTCAGGCGAGTTTTCAGAACTCCAGAAATCACCGTATTGTCAGATCGAGGCCA 707
Db 583 ACCAAAGGTCAGGCGAGTTTTCAGAACTCCAGAAATCACCGTATTGTCAGATCGAGGCCA 642
Qy 708 AACCTAACACTCTTGTTCTTCCCAAGCAGCGTGAATCTGATGAACATCCCTGTTATCCAGC 767
Db 643 GACCTAACACTCTTGTTCTTCCCAAGCAGCGTGAATCTGATGAACATCCCTGTTATCCAGC 702
Qy 768 AAGGCGAAGCCACCGTAGCGTAGCAAAATGGCAATTAACAAGAGACTTTAATCTTGAGC 827
Db 703 AAGGCAAGCCACCGTAGCGTAGCAAAATGGCAATTAACAAGAGACTTTAATCTTGAGC 762
Qy 828 AGGGCCATGCACCTCAGAAATCCCATCCGTTTCAATTCCTACATCTTGAACCCCATGACA 887
Db 763 AGGGCCATGCACCTCAGAAATCCCATCCGTTTCAATTCCTACATCTTGAATCCCATGACA 822
Qy 888 ACCAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAACACACCCGCCAGTTTGAGG 947
Db 823 ACCAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAACACACCCGCCAGTTTGAGG 882
Qy 948 ATTCTTCCGGGAGCAGCGAGACCAATCATCTTCTGACAGGCTTCAGAGGAATA 1007
Db 883 ATTCTTCCGGGAGCAGCGAGACCAATCATCTTCTGACAGGATTCAGAGGAATA 942
Qy 1008 CGTTGAGCGCCGCTTCAATGCGGAATTCATAGATACGAGGGTGCTGTTAGAGAGA 1067
Db 943 CTTTGAGCGCCGCTTCAATGCGGAATTCATAGATACGAGGGTGCTGTTAGAGAGA 1002
Qy 1068 ATGCGAGGAGTGAGCAAGAGGAGAGGCGCAGAGCGATCTCGGAGTAGTGAGA 1127
Db 1003 ATGCGAGGAGTGAGCAAGAGGAGAGGCGCAGAGCGATCTCGGAGTAGTGAGA 1059
Qy 1128 ACAATGAAGAGTGATAGTCAAGTGTCAAAGAGCACGTTGAAGAACTTACTAGACGAG 1187
Db 1060 ATAATGAAGAGTGATAGTCAAGTGTCAAAGAGCACGTTGAAGAACTTACTAGACGAG 1119
Qy 1188 CTAAATCCGCTCAAAAGAAAGGCTCGAAGAGAGAGGAGATATCAACCAACCAATCACT 1247
Db 1120 CTAAATCCGCTCAAAAGAAAGGCTCC---GAAGAGGAAGATATCAACCAACCAATCACT 1176
Qy 1248 TGAGAGAGGCGAGCGCCGATCTTCTTCAACACTTTTGGGAAGTTATTGAGGTGAAGCCAG 1307
Db 1177 TGAGAGATGGCGAGCCGATCTTCTTCAACACTTTTGGGAAGTTATTGAGGTGAAGCCAG 1236
Qy 1308 ACAAGAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAG 1367
Db 1237 ACAAGAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAG 1296
Qy 1368 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGAGCCATGTTATCGTCGTCGTCACA 1427
Db 1297 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGAGCCATGTTATCGTCGTCGTCACA 1356
Qy 1428 AAGGAACTGGAAACCTTGAACCTCGTGCTGTGAAGAAAGAGCAACAACAGAGGGGACGCG 1487
Db 1357 AAGGAACTGGAAACCTTGAACCTCGTAGCTGTGAAGAAAGAGCAACAACAGAGGGGACGCG 1416
Qy 1488 GGGAA-----GAAGAGGAGGAGCAAGAGCAAGAGAGGAGGAAGTAACAGAGAGG 1538
Db 1417 GGGAAACAGAGTGGGAAGAGAGGAGGAAGATGAAGAGAGGAGGGAAGTAACAGAGAGG 1476
Qy 1539 TCGTAGGTATACAGCAGGAGTTGAAGAGAGGAGTGTGTTCAATCATGCCAGAGCTCATC 1598
Db 1477 TCGTAGGTATACAGCAGGAGTTGAAGAGAGGAGTGTGTTCAATCATGCCAGAGCTCATC 1536
Qy 1599 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAA 1596
Qy 1659 ACAACCAAGATCTCTCTGAGGTGATGAAGCAATGTGTAGACCCAGATAGAGAAGC 1718
|||||

Db 1597 ACAACCAAGAAATCTTCTTGCAAGTGATAAGGACAATGTGTAGACCAAGATAGAGAAGC 1656
Qy 1719 AAGCCGAAGATTTAGCATTTCCCTGGGTGGGTGAACAAGTTGAGAGGCTCATCAAAAACC 1778
Db 1657 AAGCCGAAGATTTAGCATTTCCCTGGGTGGGTGAACAAGTTGAGAGGCTCATCAAAAACC 1716
Qy 1779 AGAAGAAATCTCATTGTTGTGAGTGCTGCTCAATCTCAATCTCAATCTCGTCTC 1838
Db 1717 AGAGGAGTCTCATTGTTGTGAGTGCTGCTCAATCTCAATCTCGTCTC-----1765
Qy 1839 CTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGAGGAAAAACCAAGAGGGAAGGCTC 1898
Db 1766 -----GTCTCTGAAAAAGAGGATCAAGAGAGGAAAAACCAAGAGGGAAGGCTC 1815
Qy 1899 CACTCTTTCAATTTGAAAGGCTTTTAACTGAGAAATGAGGCAACTTGTATGTATCGAT 1958
Db 1816 CACTCTTTCAATTTGAAAGGCTTTTAACTGAGAAATGAGGCAACTTGTATGTATCGAT 1875
Qy 1959 AATAAGATCAGCTTTTGTACTTACTATCCAAAACTTATCAATAAATAAAAAAGCTTTG 2018
Db 1876 AATAAGATCAGCTTTTGTAACTACTATCCAAAACTTATCAATAAATAAAAAAGCTTTG 1935
Qy 2019 TCGTGTGTTCTCC 2032
Db 1936 TCGTGTGTTCTCC 1949

RESULT 10
AAS08539
ID AAS08539 standard; cDNA; 1952 BP.
AC AAS08539;
XX
DT 23-OCT-2001 (first entry)
XX
DE Anaphylactic antigen Ara h 1, cDNA clone P17.
XX
KW Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
XX
KW allergy; mast cell; basophil; mouse; ss.
XX
OS Mus sp.
XX
PN WO200140264-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033124.
XX
PR 06-DEC-1999; 99US-00455294.
XX
PR 23-JUN-2000; 2000US-0213765P.
XX
PR 27-SEP-2000; 2000US-0235797P.
XX
XX (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
XX Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX
XX WPI; 2001-381378/40.
XX
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals sensitive
PT to antigens, have reduced ability to bind Immunoglobulin E.
XX
XX Disclosure; Fig 9; 100pp; English.
XX
CC The sequence represents anaphylactic antigen Ara h 1, cDNA clone P17. Ara
CC h 1 is an anaphylactic antigen (A), which was used to design antigenic
CC peptides having a reduced ability to bind immunoglobulin E (IgE) as
CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is

CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 1952 BP; 608 A; 456 C; 512 G; 376 T; 0 U; 0 Other;

Query Match 86.1%; Score 1749.8; DB 4; Length 1952;

Best Local Similarity 95.0%; Pred. No. 0;

Matches 1893; Conservative 0; Mismatches 42; Indels 58; Gaps 6;

Qy	48	CAATGAGAGGAGGTTTCTCCACTGATGCTGTGTCTAGGATCCTTGCTCGCTTCAG	107
Db	1	CAATGAGAGGAGGTTTCTCCACTGATGCTGTGTCTAGGATCCTTGCTCGCTTCAG	60
Qy	108	TTTCTGCAAGCGATGCCAAGTCATCACTTACCAGAGAGAGAGAGAGAGAGAGAGAG	167
Db	61	TTTCTGCAAGCGAGCCCAAGTCACC-----TTACCGAGAGAGAGAGAGAGAGAGAG	114
Qy	168	AGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGT	227
Db	115	AGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGT	174
Qy	228	CTCGCTCACCAAGCTCGAGTATGATCCTCTGTGTGTCTATGATCCTCGAGGACACACTG	287
Db	175	CTCGCTCACCAAGCTCGAGTATGATCCTCTGTGTGTCTATG-----CACACTG	223
Qy	288	GCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCGAGACT	347
Db	224	GCGCCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTTGGCGCGCAACCCGAGACT	283
Qy	348	ACGATGATGACCGCGCTCAACCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	407
Db	284	ACGATGATGACCGCGCTCAACCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	343
Qy	408	CGAGGGAGCGTGAAG	467
Db	344	CGAGGGAGCGTGAAG	403
Qy	468	GTCAATCAGCAGCCACGGAAGAAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAG	527
Db	404	GTCAATCAGCAGCCACGGAAGAAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAG	463
Qy	528	CACCAAGGTAGCCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	587
Db	464	CACCAAGGTAGCCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	523
Qy	588	GGCGGTTTAGACCCGCTACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	647
Db	524	GGCGGTTTAGACCCGCTACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	583
Qy	648	ACCAAGGTCAAGGAG	707
Db	584	ACCAAGGTCAAGGAG	643
Qy	708	AACCTAACACTCTGTGTCTTCCAGCAGCCTGATGCTGATACATCATCTTGTATCCAGC	767
Db	644	GACCTAACACTCTGTGTCTTCCAGCAGCCTGATGCTGATACATCATCTTGTATCCAGC	703
Qy	768	AAGGCGAAGCCAGCGTACCGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	827
Db	704	AAGGCGAAGCCAGCGTACCGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	763
Qy	828	AGGCGCATGCACTCAGAAATCCCATCCGTTTCTATCTTCTACATCTTGAACCGCATGACA	887

Db	764	AGGCGCATGCACTCAGAAATCCCATCCGTTTCTATTTCTTACATCTTGAATCGACATGACA	823
Qy	888	ACCAGAACTCTAGAGTAGCTTAAATCTCATGCCCTTAAACACACCCGCGCAGTTTGAGG	947
Db	824	ACCAGAACTCTAGAGTAGCTTAAATCTCATGCCCTTAAACACACCCGCGCAGTTTGAGG	883
Qy	948	ATTTCTTCCGCGGAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1007
Db	884	ATTTCTTCCGCGGAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	943
Qy	1008	CGTTGAGAGCGCGCTTCAATGCGGAATTCATATGAGATACGAGAGAGAGAGAGAGAGAG	1067
Db	944	CGTTGAGAGCGCGCTTCAATGCGGAATTCATATGAGATACGAGAGAGAGAGAGAGAGAG	1003
Qy	1068	ATGAGAGAGGTGAGCAAG	1127
Db	1004	ATGAGAGAGAGAGCAAG	1060
Qy	1128	ACAATCAAGGAGTGTAGTCAAAAGTGTCAAAAGGAGCAGCGTTGAAGAACTTACTTAAGCAG	1187
Db	1061	ATAATGAGAGAGTGTAGTCAAAAGTGTCAAAAGGAGCAGCGTTCAAGAACTTACTTAAGCAG	1120
Qy	1188	CTAAATCCCTCTCAAAAGAAAGGCTCCGAAAGAGAGAGAGAGAGATATCACCAACCCCAATCAACT	1247
Db	1121	CTAAATCCCTCTCAAAAGAAAGGCTCCGAAAGAGAGAGAGAGATATCACCAACCCCAATCAACT	1177
Qy	1248	TGAGAGAGCGGAGCCCGATCTTTCTAAACAACTTTGGGAGAGTTTATTTGAGGTGAAGCCAG	1307
Db	1178	TGAGAGATGCGGAGCCCGATCTTTCTAAACAACTTTGGGAGAGTTTATTTGAGGTGAAGCCAG	1237
Qy	1308	ACAAGAGAAACCCCGAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAG	1367
Db	1238	ACAAGAGAAACCCCGAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAG	1297
Qy	1368	AAGGAGCTTTGATGCTCCACACTTCAAAAGGCACTGGTTATTCGTCGTCTCAACA	1427
Db	1298	AAGGAGCTTTGATGCTCCACACTTCAAAAGGCACTGGTTATTCGTCGTCTCAACA	1357
Qy	1428	AAGGAACTCGGAAACCTTTGAACTCGTAGCTCGACATGATGCTCACCTGTGTAGAGATCAAG	1487
Db	1358	AAGGAACTCGGAAACCTTTGAACTCGTAGCTCGACATGATGCTCACCTGTGTAGAGATCAAG	1417
Qy	1488	GGGAA-----GAGAGGAGT	1539
Db	1418	GGGAA-----GAGAGGAGT	1477
Qy	1540	GGCTAGGTACACAGCGAGGTTGAAGGAGGCGATGTTTTCATCATGCCAGCAGCTCATCC	1599
Db	1478	GGCTAGGTACACAGCGAGGTTGAAGGAGGCGATGTTTTCATCATGCCAGCAGCTCATCC	1537
Qy	1600	AGTAGGCATCAACGCTTCTCCGAACTCATCTGTGTGGCTTCGGTATCAACGCTGAAAA	1659
Db	1538	AGTAGGCATCAACGCTTCTCCGAACTCATCTGTGTGGCTTCGGTATCAACGCTGAAAA	1597
Qy	1660	CAACACAGAACTTCTTCGAGGTGATAGGAGCAATGTGATAGACCATAGAGAGAGCA	1719
Db	1598	CAACACAGAACTTCTTCGAGGTGATAGGAGCAATGTGATAGACCATAGAGAGAGCA	1657
Qy	1720	AGCGAAGGATTTAGCAATTCCTGGGTTCGGGTGAAACAAAGTTGAGAACTCATCAAAAAACCA	1779
Db	1658	AGCGAAGGATTTAGCAATTCCTGGGTTCGGGTGAAACAAAGTTGAGAACTCATCAAAAAACCA	1717
Qy	1780	GAAGGAATCTCATTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGCTCTCTCC	1839
Db	1718	GAGGAGTCTCATTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGCTCTCTCC	1759
Qy	1840	TGAGAAAGAGTCTCTGAGAAAGAGAGATCAAGGAGGAGAGAGAGAGAGAGAGAGAGTCC	1899
Db	1760	-----GTCTCTGAGAAAGAGAGATCAAGGAGGAGAGAGAGAGAGAGAGAGAGTCC	1810
Qy	1900	ACTCTCTTCAATTTTGAAGGCTTTTAACTGAGAAAGAGAGAGAGAGAGAGAGAGAGTCCGAT	1959
Db	1811	ACTCTCTTCAATTTTGAAGGCTTTTAACTGAGAAAGAGAGAGAGAGAGAGAGAGAGTCCGAT	1870

Query Match 20.0%; Score 406.8; DB 9; Length 1254;
Best Local Similarity 62.1%; Pred. No. 3.4e-109;
Matches 779; Conservative 0; Mismatches 382; Indels 93;

Db	1014	TTTTGGTATCAATCTGAGAACCAACGAGGAACTTCCTTGCAGCGAGAAAGACAATGT	1073
Qy	1699	GATAGACCAGATAGAGAAGCAAGCGAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGT	1758
Db	1074	GGTAAAGCAGATAGAAAGACAAAGTGCAGGACCTTCGCTTCCTCGGTCCTGCACAAGATGT	1133
Qy	1759	TGAGAAGCTCATCAAAACCAGAAGGAATTCATCTTTGTGAGTGCTCGTCCTCA	1812
Db	1134	TGAGAGGCTATTAAAGAAAGCAGAGGGAATCTCACTTTGTGTATGCTCAGCCTCA	1187

RESULT 13

ACC49561

ID ACC49561 standard: cDNA: 1278 BP.

XX XX

AC ACC49561;

DT 01-JUL-2003 (first entry)

XX DE FLAG-tagged beta-conglycinin beta subunit encoding cDNA SEQ ID NO:13.

Glycine max; soybean; 7S-beta-conglycinin beta subunit; plant; food;
KW nutritional supplementation; animal feed; nutritive value; malnutrition;
KW health; gene; ss.

Glycine max.

OS
OS
OS

32 XX
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1

Item	Key	Location/Qualifiers
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

FT	CDS	1. .1278

FT /*taq= a

```
FT
/product= "FLAG-tagged beta-conglycinin beta subunit"
```

XX

PN WO2003025153-A2.

XX

PD 27-MAR-2003.

XX

PF 17-SEP-2002; 2002WO-US030373.

[illegible]

PR 17-SEP-2001; 2001US-

XX

PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX
PI Rapp WD, Peng J, Nadig G, Venkatesh T;
XX
XX
DR WPI: 2003-333163/31.
XX P-PSDB; ABP97238.
XX
PT Novel modified polypeptide useful in nutritional supplementation for
PT animals and as human food, comprises a substitution of essential amino
PT acids such as lysine, threonine and tryptophan into an unmodified
PT polypeptide.

Example 6: Page 88-89: 107pp: English.

XX The present invention describes a modified polypeptide (I) comprising a
CC substitution of one or more essential amino acids selected from
CC isoleucine, lysine, methionine, threonine and tryptophan into an
CC unmodified polypeptide having a sequence (SI) of 439 amino acids (see
CC ABP97235) of the Glycine max 7S-beta-conglycinin beta subunit, where the
CC modified polypeptide is capable of accumulating in a seed. Also
CC described: (1) a recombinant nucleic acid molecule (II) encoding (I),
CC preferably a modified beta-conglycinin polypeptide, where the modified
CC beta-conglycinin polypeptide is capable of accumulating in a cell; (2) a
CC cell (III) comprising (II); (3) a transgenic plant (IV) comprising (III);
CC a seed (V) from (IV); and (5) an animal feed comprising (V). (I) is
CC useful in nutritional supplementation for animals, including humans. (I),
CC (IV) and (V) are useful as human food. (I) significantly improves the
CC nutritive value of animal feed, leading to improved quality and quantity
CC of commercial animal products, and also significantly improves the
CC nutritive value of human food, leading to a decreased incidence of
CC malnutrition and associated health problem and improves the overall

Qy	1339	CATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCACACTTCAACTC	1398
Db	813	TATCTTCTCTCAGTTCTGTGGATATCAACGAGGAGCTCTTCTTCTACACACTTCAATTC	872
Qy	1399	AAAGCCATCGTTATCGTCGTCGTCACAAGAGAACTGGAAAACTTCGAGCTCGTGGCTGT	1458
Db	873	AAAGCCATAGTGATACTAGTGATTATGAAGGAGATGCAAACTTGGACTTTGTTGGCAT	932
Qy	1459	AAGAAAAGAGCAACAA CAGAGGGGACGGCGGGAAAGAGAGGAGCAAGAACGGAAGAAGA	1518
Db	933	TAAAGAAACAA CAA CAGAGAGCAGAAACAGGAAGAGGAACTTTG-----	975
Qy	1519	GGAGGGAATACAGAGAGGTGCGTAGGTACACAGCGAGGTTCGAAGAAAGCGATGTGTT	1578
Db	976	-----GAAGTGCAAAGGTACAGAGCTGAATTTGCTGAAGACGATGATT	1019
Qy	1579	CATCATCCAGCAGCTCATCCAGTAGGCATCAACGCTTCCTCCGAACTCCACTCTGCTGG	1638
Db	1020	TGTAATTCACGACGCTTATCCATTTGTGTCGTCACGCTACCTCAAACTCAATTTCCCTTCG	1079
Qy	1639	CTTCGGTATCAACGCTGAAAACACCAACAGATCTTCTTCGTGAGGTGATTAAGACAAATGT	1698
Db	1080	TTTGTGTTCAATGCTGAGAACCAACAGAGGAATCTCTTCGAGGCGGAGAAAGACAAATGT	1139
Qy	1699	GATAGACAGATAGAGAAGCAAGCGAAGGATTTTAGCATTCCTCTGGGTCGGTCAACAAGT	1758
Db	1140	GGTAAAGCAGATAGAAGAACAAGTGCAGGAGCTTCGCTTCCTGGGTCTGCACAAAGTGT	1199
Qy	1759	TGAGAAGCTCATCAAAACCAGAAAGGAATCTCACTTTGTGAGTGCTCGTCTCA	1812
Db	1200	TGAGAGGCTATTAAAGAAAGCAGAGGGAATCTACTTTGTGTGCTCAGGCTCA	1253

Search completed: August 23, 2005, 21:26:04
Job time : 773.871 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 334.307 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aatacatatataatcatc.....cgtttgctgctgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	4	US-09-106-872A-21
2	2032	100.0	2032	4	US-09-191-593-5
3	2032	100.0	2041	4	US-09-106-872A-23
4	1930	95.0	1930	4	US-09-106-872A-3
5	1752.4	86.2	1949	4	US-09-106-872A-19
6	1752.4	86.2	1949	4	US-09-191-593-4
7	750	36.9	750	4	US-09-106-872A-18
8	406.8	20.0	1320	3	US-09-108-010B-3
9	406.8	20.0	1320	4	US-09-758-552-3
10	406.8	20.0	1320	4	US-10-684-651-3
11	404	19.9	1818	3	US-09-108-010B-1
12	404	19.9	1818	4	US-09-758-552-1
13	404	19.9	1818	4	US-10-684-651-1
14	324.2	16.0	1920	3	US-09-108-010B-2
15	324.2	16.0	1920	4	US-09-758-552-2
16	324.2	16.0	1920	4	US-10-684-651-2
17	138.8	6.8	1867	1	US-07-955-905A-1
18	98.8	4.9	1924	3	US-09-424-283-5
19	71.6	3.5	7218	1	US-08-232-463-14
20	49.4	2.4	1068	4	US-09-710-794-3
21	46.6	2.3	74177	4	US-09-949-016-11988
22	46.6	2.3	74177	4	US-09-949-016-17388
23	46.6	2.3	1590	3	US-09-323-195A-7
24	45.8	2.3	300402	4	US-09-949-016-13632
25	44.4	2.2	499	4	US-10-235-618B-16
26	44.2	2.2	61664	4	US-09-949-016-13308
27	43.8	2.2	767677	4	US-09-949-016-12147

28	43.8	2.2	767677	4	US-09-949-016-17361	Sequence 17361, A
29	43.4	2.1	1318	4	US-10-235-618B-1	Sequence 1, Appl
30	43.4	2.1	1318	4	US-10-235-618B-11	Sequence 11, Appl
31	43.2	2.1	289	3	US-09-007-005-17	Sequence 17, Appl
32	43.2	2.1	289	3	US-09-244-796-17	Sequence 17, Appl
C 33	42	2.1	150394	4	US-09-949-016-13042	Sequence 13042, A
C 34	41.8	2.1	28806	4	US-09-949-016-13217	Sequence 13217, A
35	41	2.0	73519	4	US-09-949-016-16344	Sequence 16344, A
36	41	2.0	105919	4	US-09-949-016-11769	Sequence 11769, A
C 37	40.8	2.0	50383	4	US-09-949-016-17600	Sequence 17600, A
C 38	40.8	2.0	234884	4	US-09-949-016-16420	Sequence 16420, A
C 39	40.6	2.0	46823	4	US-09-949-016-12723	Sequence 12723, A
C 40	40.6	2.0	46940	4	US-09-949-016-16252	Sequence 16252, A
C 41	40.6	2.0	69813	4	US-09-949-016-12455	Sequence 12455, A
C 42	40.6	2.0	69813	4	US-09-949-016-13906	Sequence 13906, A
C 43	40.6	2.0	69833	4	US-09-949-016-12861	Sequence 12861, A
C 44	40.6	2.0	69833	4	US-09-949-016-12861	Sequence 12861, A
C 45	40.4	2.0	601	4	US-09-949-016-167942	Sequence 167942, A

ALIGNMENTS

RESULT 1
US-09-106-872A-21
; Sequence 21, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106, 872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-106-872A-21

Query Match	100.0%;	Score 2032;	DB 4;	Length 2032;
Best Local Similarity	100.0%;	Pred. No. 0;		
Mismatches 2032;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Db	1	AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Qy	61	GGTTTCTCCACTGATGCTGTGTGCTAGGGATCTTGTCTGCTGCTTTCAGTTTCTGCAAGCA	120	
Db	61	GGTTTCTCCACTGATGCTGTGTGCTAGGGATCTTGTCTGCTGCTTTCAGTTTCTGCAAGCA	120	
Qy	121	TGCAAGTATCATCTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCTTCCA	180	
Db	121	TGCAAGTATCATCTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCTTCCA	180	
Qy	181	GAGTTCTCAACAGAACCCGATGACCTTGAAGCAAAAGGATCGAGTCTCGCTGCACCA	240	
Db	181	GAGTTCTCAACAGAACCCGATGACCTTGAAGCAAAAGGATCGAGTCTCGCTGCACCA	240	
Qy	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA	300	
Db	241	GCTCGAGTATGATCCTCGTTGTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA	300	

Qy 301 ACCTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Db 301 ACCTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Qy 361 CCCTCAACCCCGAAG 420
Db 361 CCCTCAACCCCGAAG 420
Qy 421 AAGAGAAAGAGACTGGAGACAAACCAAGAGAGATTTGGAGCGCACCAAGTCAACAGAGCC 480
Db 421 AAGAGAAAGAGACTGGAGACAAACCAAGAGAGATTTGGAGCGCACCAAGTCAACAGAGCC 480
Qy 481 ACAGAAATAGGCCCCGAAG 540
Db 481 ACAGAAATAGGCCCCGAAG 540
Qy 541 TGTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTCCCGTCAAGGCGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTCCCGTCAAGGCGTTTAGCAC 600
Qy 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGCTCTGCAGAGGTTTGAACCAAGGTTCAAG 660
Db 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGCTCTGCAGAGGTTTGAACCAAGGTTCAAG 660
Qy 661 GCAGTTTTCAGAACTCCAGAAATCACCGTATTTGTCAGATCGAGGCCAAACCTTAACACTCT 720
Db 661 GCAGTTTTCAGAACTCCAGAAATCACCGTATTTGTCAGATCGAGGCCAAACCTTAACACTCT 720
Qy 721 TGTTCCTCCAGACGCTGATGCTGATTAACATCTCTGTTATCGAGCAAGGGCAAGCCAC 780
Db 721 TGTTCCTCCAGACGCTGATGCTGATTAACATCTCTGTTATCGAGCAAGGGCAAGCCAC 780
Qy 781 CGTGACCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAACTCTTGACGAGGSCCATGCAC 840
Db 781 CGTGACCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAACTCTTGACGAGGSCCATGCAC 840
Qy 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGCCATGACAAACAGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGCCATGACAAACAGAACTCAG 900
Qy 901 AGTAGCTAAATCTCCATGCGGTTTAAACACACCGGCGCAGTTTGAGGATTTCTTCCGGC 960
Db 901 AGTAGCTAAATCTCCATGCGGTTTAAACACACCGGCGCAGTTTGAGGATTTCTTCCGGC 960
Qy 961 GAGCAGCGGAGACCAATCATCTTCTACTTGAGGGGCTTCAGCAGGAATACGTTGGAGCGCG 1020
Db 961 GAGCAGCGGAGACCAATCATCTTCTACTTGAGGGGCTTCAGCAGGAATACGTTGGAGCGCG 1020
Qy 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGCTGTTAGAGAGAGATGAGAGAGAGTGA 1080
Db 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGCTGTTAGAGAGAGATGAGAGAGAGTGA 1080
Qy 1081 GCAAGAGGAGAGGGCAGAGGCGATGAGTACTCGGAGTAGTGAGAACATCAAGAGGT 1140
Db 1081 GCAAGAGGAGAGGGCAGAGGCGATGAGTACTCGGAGTAGTGAGAACATCAAGAGGT 1140
Qy 1141 GATAGTCAAAAGTGTCAAAGGAGCAGCTGTGAAGAACTTACTAAGCAGCCTAAATCCGCTCT 1200
Db 1141 GATAGTCAAAAGTGTCAAAGGAGCAGCTGTGAAGAACTTACTAAGCAGCCTAAATCCGCTCT 1200
Qy 1201 AAAGAAAGGCTCCGAAGAGAGGGAGATATCAACAAACCCCAATCAACTTTGAGAGAGCGGA 1260
Db 1201 AAAGAAAGGCTCCGAAGAGAGGGAGATATCAACAAACCCCAATCAACTTTGAGAGAGCGGA 1260
Qy 1261 GCCCGATCTTTTAAACACTTTGGGAGTTATTTGAGGTGAAGCCAGACAGAAAGAACCC 1320
Db 1261 GCCCGATCTTTTAAACACTTTGGGAGTTATTTGAGGTGAAGCCAGACAGAAAGAACCC 1320
Qy 1321 CCAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTGAT 1380
Qy 1381 GCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCGTCTCAACAAAGGAACTGGAAA 1440

Db 1381 GCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCGTCTCAACAAAGGAACTGGAAA 1440
Qy 1441 CTTTGAACCTCGCTGCTGAAGAAAGAGCAACAAACAGAGGGGACCGCGGGAAGAGAGGA 1500
Db 1441 CTTTGAACCTCGCTGCTGAAGAAAGAGCAACAAACAGAGGGGACCGCGGGAAGAGAGGA 1500
Qy 1501 GGACGAAGACGGAAGTT 1560
Db 1501 GGACGAAGACGGAAGTT 1560
Qy 1561 GAAGGAAGCGGATGTTTCAATCATGTCAGAGCTCATCCAGTAGCCTCAACGCTTCTCTC 1620
Db 1561 GAAGGAAGCGGATGTTTCAATCATGTCAGAGCTCATCCAGTAGCCTCAACGCTTCTCTC 1620
Qy 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAAATCTTCTCTGC 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAAATCTTCTCTGC 1680
Qy 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTTCCC 1740
Db 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTTCCC 1740
Qy 1741 TGGGTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGCTCGCTCGCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGCTCGCTCGCTGAGAAAGAGTCTCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Db 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Qy 1921 TTTTAACTGAGAAATGAGGCAACTTGTATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 TTTTAACTGAGAAATGAGGCAACTTGTATGATGATGATGATGATGATGATGATGATGATGAT 1980
Qy 1981 CTACTATCCAAACCTTATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2032
Db 1981 CTACTATCCAAACCTTATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2032

RESULT 2
US-09-191-593-5
; Sequence 5, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/717,933
FILING DATE: 23 SEPTEMBER 1996
APPLICATION NUMBER: US 07/998,377
FILING DATE: 30 DECEMBER 1992
APPLICATION NUMBER: US 08/158,704
FILING DATE: 29 NOVEMBER 1993
APPLICATION NUMBER: US 60/009,455
FILING DATE: 29 DECEMBER 1995
APPLICATION NUMBER: US 08/610,424
FILING DATE: 04 MARCH 1996
ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2032 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: identified as Ara h I Beta P41b
HYPOTHETICAL: No
ANTI-SENSE: No
FRAGMENT TYPE: No. 6835824 applicable
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE: Clone P41b
DEVELOPMENTAL STAGE: Seed
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE: Seed mRNA, cDNA library
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY: Florunner seed cDNA expression
LIBRARY: library in Uni-ZAP XR vector
CLONE: P41b
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY: CDS
LOCATION: 50..1930
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-09-191-593-5

Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATAGAGGGAG	60
Db	1		
Qy	61	GGTTTCTCCACTGATGCTGTGGTAGGGATCCTTGTCTCGCTTCAGTTCTGCAACGCA	120
Db	61		
Qy	121	TGCCAAGTCATCACTTACCCAGAGAAACAGAGAACCCCTCGGCCAGAGGTGCCTCCA	180
Db	121		
Qy	181	GAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCA	240
Db	181		
Qy	241	GCTCGAGTATGATCCTCGTTGTGTATGATCTCGAGGACACACTGGCACCAACCA	300
Db	241		
Qy	301	ACGTTCCCTCCAGGGAGCGGACACGTGCCGCCCAACCCGGAGACTACGATGATACCG	360
Db	301		
Qy	361	CCGTCAACCCGAGAGAGAGAGGCGCGATGGGACCGAGCTGGACCGGAGGCGTGA	420
Db	361		
Qy	421	AAGAGAAAGAGACTGGAGACCAACAGAGAAGATTGGAGCGCACCAAGTCAATCAGCAGCC	480
Db	421		
Qy	481	ACGAAAAATAAGGCCCGGAGAGAGAGAGAACAGAGTGGGGAAACACCAAGTAGCCA	540
Db	481		
Qy	541	TGTGAGGAGAAACATCTCGGAAACAAACCTTTCTACTTCCCGTCAAGGGCGTTAGCAC	600
Db	541		
Qy	601	CCCTACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTCACCAAGGTCAG	660
Db	601		
Qy	661	GCAGTTTCAGAACTCCAGAAATCACCGTATTTGTGCAGATCGAGGCCAAACCTAACACTCT	720
Db	661		
Qy	721	TGTTCTTCCCAAGCAGCTGATGCTGATACATCTTGTATCCAGCAAGGCGCAAGCCAC	780
Db	721		
Qy	781	CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCATGCAC	840
Db	781		
Qy	841	CAGAAATCCCATCCCGTTTCATTTCTTACATCTTTGAACCGCCATGACAAACAGAACCTCAG	900
Db	841		
Qy	901	AGTAGCTAAATCTCCATGCCCGCTTAAACACACCCGGCGAGTTTGAGGATTTCTTCCCGGC	960
Db	901		
Qy	961	GAGCAGCCGAGACCAATCATCTTCTTGTGAGGGCTTCAGCAGGAATACGTTGGAGGCCGC	1020
Db	961		
Qy	1021	CTTCAATGCGGAATTCAATGAGATACGAGGGTGTCTGTTAGAGAGAAATCAGGAGGTGA	1080
Db	1021		
Qy	1081	GCAAGAGGAGAGAGGGCAGAGGCGATGGAGTACTCGGAGTAGTAGAGAACAAAGAGAGT	1140

Db 1081 GCAAGAGGAGAGAGGCGAGAGGCGATGAGTACTCGGAGTAGTGAGAAACAATGAAGAGT 1140
Qy 1141 GATAGTCAAAAGTGTCAAGAGGACAGTGTGAAGAACTTACTTAAGCAGCGTAAATCCGCTTC 1200
Db 1141 GATAGTCAAAAGTGTCAAGAGGACAGTGTGAAGAACTTACTTAAGCAGCGTAAATCCGCTTC 1200
Qy 1201 AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTTGAGAGAAGCGGA 1260
Db 1201 AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTTGAGAGAAGCGGA 1260
Qy 1261 GCCGATCTTTCTAAACAACCTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCC 1320
Db 1261 GCCGATCTTTCTAAACAACCTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCC 1320
Qy 1321 CCAGCTTCAGGACTGGACATGATGCTCACCTGTGTAGAGTCAAGAGGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACTGGACATGATGCTCACCTGTGTAGAGTCAAGAGGAGCTTTGAT 1380
Qy 1381 GCTCCACACACTTCAACTCAAAAGCCATGTTATCGTGTGCTCAACAAAGGAACTCGAAA 1440
Db 1381 GCTCCACACACTTCAACTCAAAAGCCATGTTATCGTGTGCTCAACAAAGGAACTCGAAA 1440
Qy 1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGGGA 1500
Db 1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGGGA 1500
Qy 1501 GAGCAGAGCAAGAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGCTACACAGCGAGGTT 1560
Db 1501 GAGCAGAGCAAGAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGCTACACAGCGAGGTT 1560
Qy 1561 GAAGGAAGGCGATGTGTTCAATCATGCCAGAGCTCATCCAGTAGCCATCAACCGCTTCCCTC 1620
Db 1561 GAAGGAAGGCGATGTGTTCAATCATGCCAGAGCTCATCCAGTAGCCATCAACCGCTTCCCTC 1620
Qy 1621 CGAATCTCAATCTGCTGGCTTCGGTATCAACCGCTGAAACCAACAGAGATCTTCCCTTGC 1680
Db 1621 CGAATCTCAATCTGCTGGCTTCGGTATCAACCGCTGAAACCAACAGAGATCTTCCCTTGC 1680
Qy 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGGAGCGAAGGATTTAGCATTTCCC 1740
Db 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGGAGCGAAGGATTTAGCATTTCCC 1740
Qy 1741 TGGGTCCGGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCCGGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCCTGAGAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTCTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTGAGAATGGAGGCAACTTGTATGATATCGATATAAGATCACGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAATGGAGGCAACTTGTATGATATCGATATAAGATCACGCTTTTGTACT 1980
Qy 1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGCTTTGTGCGTGTGTTCTCC 2032
Db 1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGCTTTGTGCGTGTGTTCTCC 2032

RESULT 3

US-09-106-872A-23
; Sequence 23, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-106-872A-23

Query Match 100.0%; Score 2032; DB 4; Length 2041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGGAG 60
Qy 61 GGTTTCTCCACTGATGCTGTGCTAGGGATCCTTGTCTGGCTTCAAGTTTCTGCAACGCA 120
Db 61 GGTTTCTCCACTGATGCTGTGCTAGGGATCCTTGTCTGGCTTCAAGTTTCTGCAACGCA 120
Qy 121 TGCAAAGTCATCACCTTACAGAGAAACACAGAGAAACCCCTGGCCAGAGGTGCCTCCA 180
Db 121 TGCAAAGTCATCACCTTACAGAGAAACACAGAGAAACCCCTGGCCAGAGGTGCCTCCA 180
Qy 181 GAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGATCGAGTCTCGCTGCAACCAA 240
Db 181 GAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGATCGAGTCTCGCTGCAACCAA 240
Qy 241 GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCAACCAACCA 300
Db 241 GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCAACCAACCA 300
Qy 301 AGTTTCCCTCCAGGGAGCGGACACGTTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Db 301 AGTTTCCCTCCAGGGAGCGGACACGTTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Qy 361 CGGTCAACCCCGAAGAGAGAGAGGAGGCGGATGGGGACCCAGCTGGACCCAGGAGGCGTGA 420
Db 361 CGGTCAACCCCGAAGAGAGAGAGGAGGCGGATGGGGACCCAGCTGGACCCAGGAGGCGTGA 420
Qy 421 AAGAGAAGAGAGTGGAGACAACCAAGAGAAGATTGGAGCGGACCAAGTCAATCAGAGCC 480
Db 421 AAGAGAAGAGAGTGGAGACAACCAAGAGAAGATTGGAGCGGACCAAGTCAATCAGAGCC 480
Qy 481 ACGGAAAATAAGGCCCGAAGAGAGAGAGAGAGACAAGAGTGGGGAACCAACAGGTAGCCA 540
Db 481 ACGGAAAATAAGGCCCGAAGAGAGAGAGAGAGACAAGAGTGGGGAACCAACAGGTAGCCA 540
Qy 541 TGTGAGGGAAGAAAACATCTCGGAAACAACTTTCTACTTCCCGTCAAGCGGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAAACATCTCGGAAACAACTTTCTACTTCCCGTCAAGCGGTTTAGCAC 600
Qy 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCGAGAGGTTTGAACAAAGGTCAAG 660
Db 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCGAGAGGTTTGAACAAAGGTCAAG 660
Qy 661 GCAGTTTTCAGAAATCTCAGAAATCACCGTATTGTGAGATCGAGGCCAAACCTTAACACTCT 720
Db 661 GCAGTTTTCAGAAATCTCAGAAATCACCGTATTGTGAGATCGAGGCCAAACCTTAACACTCT 720
Qy 721 TGTTCCTTCCCAAGCAGCTGATGATCAATCACTTGTGTTATCCAGCAAGGCGCAAGCCAC 780
Db 721 TGTTCCTTCCCAAGCAGCTGATGATCACTCTTGTGTTATCCAGCAAGGCGCAAGCCAC 780


```
QY 1068 ATGAGGAGGTGAGCAAGAGGAGAGAGCGGCAGAGCGATGAGTACTCGGAGTAGTGAGA 1127
Db 1003 ATGAGGAGGAGAGCAAGAGGAGAGAGAGCGGCAGAGCGATGAGTACTCGGAGTAGTG 1059
QY 1128 ACAATGAAGAGGTGATAGTCAAAGGTGCAAGAGGAGCAGCTTTGAAGAACTTACTTAAGCACG 1187
Db 1060 ATATGAAGAGGTGATAGTCAAAGGTGCAAGAGGAGCAGCTTCNAGAACTTACTTAAGCACG 1119
QY 1188 CTAATCCGTCTCAAAGAAAGGCTCCGAGAGGAGGAGATATACCAACCAATCAACT 1247
Db 1120 CTAATCCGTCTCAAAGAAAGGCTCC---GAAGAGGAAGATATACCAACCAATCAACT 1176
QY 1248 TGAGAGAGGCGAGCGCCGATCTTTCTAACAACTTTGGGAAGTATTGAGGTGAAGCCAG 1307
Db 1177 TGAGAGATGGCGAGCGCCGATCTTTCTAACAACTTTGGGAGGTATTGAGGTGAAGCCAG 1236
QY 1308 ACAAGAAAGACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAG 1367
Db 1237 ACAAGAAAGACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAG 1296
QY 1368 AAGGAGCTTTGATGCTCCCACTTCAACTCAAAAGGCCATGTTATCGTCTGCTCAACA 1427
Db 1297 AAGGAGCTTTGATGCTCCCACTTCAACTCAAAAGGCCATGTTATCGTCTGCTCAACA 1356
QY 1428 AAGGAACTGGAAACCTTGAACCTGCTGCTGTAAGAAAGAGCAACAAGAGGCGGC 1487
Db 1357 AAGGAACTGGAAACCTTGAACCTGCTGCTGTAAGAAAGAGCAACAAGAGGCGGC 1416
QY 1488 GGGAA-----GAAGAGGAGGACGAAGACGAGGAGGAGGAGGAGTAAAGAGAGG 1538
Db 1417 GGGAAAGAGGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1476
QY 1539 TGCCTAGGTACACAGCGAGGTGAAGGAGGCGATGTTTCATCATGCCAGCAGCTCATC 1598
Db 1477 TGCCTAGGTACACAGCGAGGTGAAGGAGGCGATGTTTCATCATGCCAGCAGCTCATC 1536
QY 1599 CAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGCTTGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGCTTGGTATCAACGCTGAAA 1596
QY 1659 ACAACACAGAAATCTTCTTCAGGTGATAGGACAAATGTATAGACCAGATAGAGAGC 1718
Db 1597 ACAACACAGAAATCTTCTTCAGGTGATAGGACAAATGTATAGACCAGATAGAGAGC 1656
QY 1719 AAGCGAAGGATTTAGCAATTCCTCGGTCGGGTGAACCAAGTTGAGAAGCTCATCAAAAAC 1778
Db 1657 AAGCGAAGGATTTAGCAATTCCTCGGTCGGGTGAACCAAGTTGAGAAGCTCATCAAAAAC 1716
QY 1779 AGAAGGAATCTACCTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGCTC 1838
Db 1717 AGAAGGAATCTACCTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGCTC 1765
QY 1839 CTGAGAAAGAGTCTCCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGTC 1898
Db 1766 -----GTCTCTGAAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGTC 1815
QY 1899 CACTCTTTCAATTTGAAGCTTTTAACTCAGAAATGGAGCAACTTGTATGATCGAT 1958
Db 1816 CACTCTTTCAATTTGAAGCTTTTAACTCAGAAATGGAGGAGGAGGAGGAGGAGGTC 1875
QY 1959 AATAAGATCAGCTTTTGTACTCTACTATCAAAAACTTATCAATAAATAAAAAAGTTTG 2018
Db 1876 AATAAGATCAGCTTTTGTACTCTACTATCAAAAACTTATCAATAAATAAAAAAGTTTG 1935
QY 2019 TGCCTGTTTCTCC 2032
Db 1936 TGCCTGTTTCTCC 1949
```

RESULT 7

US-09-106-872A-18

; Sequence 18, Application US/09106872A

```
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-09-106-872A-18

Query Match 36.9%; Score 750; DB 4; Length 750;
Best Local Similarity 100.0%; Pred. No. 1.3e-218; Indels 0; Gaps 0;
Matches 750; Conservative 0; Mismatches 0;

QY 687 GTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTGTCTTCCCAAGCAGCTGATGCTG 746
Db 1 GTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTGTCTTCCCAAGCAGCTGATGCTG 60
QY 747 ATAAATCATCTTGTATTATTCAGCAAGGCGAAGCCAGCTGACCGTAGCAAAATGGCAATAACA 806
Db 61 ATAAATCATCTTGTATTATTCAGCAAGGCGAAGCCAGCTGACCGTAGCAAAATGGCAATAACA 120
QY 807 GAAAGAGCTTTAATCTTTGACGAGGCGCATGACACTCAGAAATCCCATCCGGTTTCATTTCCT 866
Db 121 GAAAGAGCTTTAATCTTTGACGAGGCGCATGACACTCAGAAATCCCATCCGGTTTCATTTCCT 180
QY 867 ACATCTTTGAACCGCCATGACAAACAGAACTCTCAGAGTAGCTAAATCTCCATGCCCGTTA 926
Db 181 ACATCTTTGAACCGCCATGACAAACAGAACTCTCAGAGTAGCTAAATCTCCATGCCCGTTA 240
QY 927 ACACACCGGCGAGTTTGAGGATTTCTTCCCGGAGCAGCGAGAGCAATCATCTACT 986
Db 241 ACACACCGGCGAGTTTGAGGATTTCTTCCCGGAGCAGCGAGAGCAATCATCTACT 300
QY 987 TGCAGGCGCTTCAGCAGGAATACGTTGGAGGCGCGCTTCAATGCGGAATTCATGAGATAC 1046
Db 301 TGCAGGCGCTTCAGCAGGAATACGTTGGAGGCGCGCTTCAATGCGGAATTCATGAGATAC 360
QY 1047 GGAGGGTGTCTTTAGAGAGAAATGCGAGGAGTTCAGCAAGAGGAGAGAGGCGAGAGGCGAT 1106
Db 361 GGAGGGTGTCTTTAGAGAGAAATGCGAGGAGTTCAGCAAGAGGAGAGAGGCGAGAGGCGAT 420
QY 1107 GGAGTACTCGGAGTAGTGAGAAACAATGAAGAGGATGATAGTCAAAGTGTCAAAGAGGACG 1166
Db 421 GGAGTACTCGGAGTAGTGAGAAACAATGAAGAGGATGATAGTCAAAGTGTCAAAGAGGACG 480
QY 1167 TTGAAGAACTTACTTAAGCAGCTAAATCCGCTCTCAAGAAAGGCTCCGAGAGAGGAGGAG 1226
Db 481 TTGAAGAACTTACTTAAGCAGCTAAATCCGCTCTCAAGAAAGGCTCCGAGAGAGGAGGAG 540
QY 1227 ATATCACCAACCAATCAACTTTGAGAGAGGCGAGCGCGATCTTTCTTAAACAACTTTGGGA 1286
Db 541 ATATCACCAACCAATCAACTTTGAGAGAGGCGAGCGCGATCTTTCTTAAACAACTTTGGGA 600
QY 1287 AGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCCAGCTTCAGGACCTGGACATGATGC 1346
Db 601 AGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCCAGCTTCAGGACCTGGACATGATGC 660
QY 1347 TCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCCACTTCAACTCAATCAAGGCCA 1406
Db 1347 TCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCCACTTCAACTCAATCAAGGCCA 1406
```

Db 661 TCACCTGTGTAGAGTCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCAAAAGGCCA 720

Qy 1407 TGGTTATCGTCGTCTCAACAAAGGAAGT 1436

Db 721 TGGTTATCGTCGTCTCAACAAAGGAAGT 750

RESULT 8

US-108-0108-3

; Sequence 3, Application US/09108010B

; Patent No. 6362399

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WINDOWS 95

; SOFTWARE: MICROSOFT WORD 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/108,010B

; FILING DATE: 30-Jun-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/019,940

; FILING DATE: JUNE 14, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: LYNN M. CRISTENBURY

; REGISTRATION NUMBER: 30,971

; REFERENCE/DOCKET NUMBER: BB-1071-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-5481

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1320 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-108-0108-3

Query Match 20.0%; Score 406.8; DB 3; Length 1320;

Best Local Similarity 62.1%; Pred. No. 1.9e-113;

Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;

Qy 562 GAACAAACCTTTCTACTTT---CCGTCGAAGCGGTTTAGCACCGCTACGCGGAACCAAAA 618

Db 90 GAATAACCCCTTCTACTTTAGAGCTCTAACAGCTTCCAAACTCTCTTTGAGAACCACAAA 149

Qy 619 CGGTAGATCGGGTCTCGAGAGGTTTGACCAAGGTCAGGAGGTTTCAGATCTCCA 678

Db 150 CGTTCCGATTCGTCTCTCCAGAGATTCAACAAACGCTCCCACTTCCCAACCTTCG 209

Qy 679 GAATCACCGGATTCGTCTCTCCAGAGTTCAACAAACGCTCCCACTTCCCAACGCGC 738

Db 210 AGACTACCGGATTCGTCTCTCCAGAGTTCAACAAACGCTCCCACTTCCCAACGCGC 269

Qy 739 TGATGCTGATAACATCTCTTTGTTATCCAGAGGCGCAAGCCGCTGACCGTAGCAATGG 798

Db 270 TGACCGCGATTTCCCTCTCTTTGTTATCCAGAGGCGCAAGCCGCTGACCGTAGCAATGG 329

Qy 799 CAATAACAGAAAGAGCTTTAAATCTTGACGAGGGCCATGCACTCAGAAATCCCATCCGGTTT 858

Db 330 CGACGACAGAGACTCTCTACAACTTTCACCTGCGATGCCAGAGAAATCCAGCTGGAAC 389

Qy 859 CATTTCCTACATCTTGAACCGCCATGACAAACCAAGAACTCTCAGAGTAGCTAAATATCTCCAT 918

Db 390 CACTTACTATTGTTTAAACCTTCAACGACCAAGAAATCTCAAAATTAATCAAACTTGCAT 449

Qy 919 GCCGTTAACACACCGCGCCAGTTTGAGGATTTCTTCCGCGGAGCAGCGAGACCAATC 978

Db 450 ACCCGTCAACAAACCTTGGCAGATATGATGATTTCTTCTTATCTAGCACTCAAGCCCAACA 509

Qy 979 ATCTTACTTTCAGGGCTTTCAGCAGGAATAGCTTGGAGGCGCCTTCAATCCGGAATTCAA 1038

Db 510 GTCTTACTTTCAGAGGCTTTCAGCAGTAATATCTAGAGACTCTCTTCTATAGCAATTCGA 569

Qy 1039 TGAGATACGAGGGTCTGTTTGAAGAGAAATGACGAGGTGACAAAGAGAGAGAGGCA 1098

Db 570 GGAGATAAAACAGGGTTTGTGAGAGGAAGAGGAGCAGAGGAGCAGCAAGAG----- 621

Qy 1099 GAGGCGATGGAGTACTCGGAGTAGTCAGAACAAATGAAGGAGTATAGTCAAGTGTCAAA 1158

Db 622 -----GGAGTGTATCGTGAAGTCTCAAA 644

Qy 1159 GGAGCAGCTTGAAGAACTTACTAAGCACGCTAAATCCGTCTCAAGAAAGGCTCCGAAGA 1218

Db 645 GGAACAAATTCGCAACTGAGCAGAGCTGCCAATCTAGTTCAAGGAACCAATTTCTCTC 704

Qy 1219 AGAGGAGAGATATCACCACCCCAATCAACTTGAAGAGGCGAGCCCGATCTTTCTAACAA 1278

Db 705 CGAAGATG-----AACCATTCAACTTGAAGCGCAACCCCATCTTATTTCCAACAA 755

Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCAGACAAAGAAAGAAACCCCGAGCTTCAGGACCTGGA 1338

Db 756 CTTTGGGAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGGACTTGA 812

Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGAGCTTTGATGCTCCACACTTCAACTC 1398

Db 813 TATCTCTCTCAGTTCTGTGATATCAACGAAGAGCTCTTCTTACCACACTTCAATTC 872

Qy 1399 AAAGGCCATGTTTATCGTCTCAACAAAGAACTGGAACCTTGAACCTCTGGGTGT 1458

Db 873 AAAGGCCATGATGATGATGATTAATGAAGGAGATGCAAAACATTTGAACCTTGTGGCAT 932

Qy 1459 AAAAAAGAGCAACACAGAGGCGGCGGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518

Db 933 TAAAGAAACAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975

Qy 1519 GGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGAGGCGATGTGT 1578

Db 976 -----GAGTGAAGGTACAGAGCTGAAATGCTCTGAGAGACGATGATT 1019

Qy 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGTTGG 1638

Db 1020 TGTAAATTCAGCAGCTTATCCATTTGCTGTCACGCTACCTCAAACTCAATTTCTCTGTC 1079

Qy 1639 CTTGCGTATCAACGCTGAAACCAACACAGAACTCTTCTTGCAGGTGATAAGGACAATGT 1698

Db 1080 TTTTGTGATCAATGCTGAGAACCAACAGAGGAACTTCTTTCAGGCGAGAGAACAAATGT 1139

Qy 1699 GATAGACCATAGATAGAGAAAGCAAGGATTTAGCAATTCCTTGGGTGCGGTGAACAAAGT 1758

Db 1140 GGTAAAGCAGATAGAGAAAGCAAGTGCAGGAGCTTGGTTCCTTGGGTCTGCACAGATGT 1199

Qy 1759 TGAGAGCTCATCAAAAAACAGAAAGAAATCTCACTTTGTGAGTGTCTGCTCTCA 1812

Db 1200 TGAGAGGCTATTAAAGAGCAGAGGGAATCTACTTTGTTGATGCTCAGCCTCA 1253

Patent No. 6703544
GENERAL INFORMATION:
APPLICANT: ANTHONY JOHN KINNEY
GARY MICHAEL FADER
TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
CLASSES OF SOYBEAN SEED
PROTEIN GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/758,652
FILING DATE: 11-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14, 1996
ATTORNEY/AGENT INFORMATION:
NAME: LYNN M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-758-652-3

Query Match 20.0%; Score 406.8; DB 4; Length 1320;
Best Local Similarity 62.1%; Pred. No. 1.9e-113;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
562 GAACAACCTTTCTACTT---CCCGTCAAGCGGTTTAGCACCCTGCTACCGGACCAAAA 618
90 GAATAACCTTTCTACTTTAGAACCTTAAAGCTTCAAGCTTCCAACTCTCTTTGAGAACCAAA 149
619 CGGTAGGATCCGGTCTCTCGAGAGGTTTGACCAAGGTCAAGGAGTTTCAGAACTCTCCA 678
150 CGTTCCGATTGCTCTCTCCAGAGATTCAACAAAGCTCCCACTTCTTTGAGAACCTTCG 209
679 GAATCAGGATTTGTCAGATCGAGGCAACCTAACTCTTTGTTCTTCCCAAGCAGC 738
210 AGACTACCGGATTTCCAGTTTTCAGTCAAAACCAACCAACAACTCTTCTCCCAACATGC 269
739 TGATGCTGATAACATCTTGTATCCAGAGGCGGACCGCTGACCGTAGCAATGG 798
270 TGACCGGATTTCTCTCTTTGCTTTAGCGGAGAGCCATCTTACCTTTGGTGAACAA 329
799 CAATAACAGAAAGAGCTTTTAACTTTGACGAGGCGCATCTCAGAACTCCCATCCGGTTT 858
330 CGACGACAGAGACTCTTACAACTTCACTTGGCGATGCCAGAGATCCAGCTGGAAC 389
859 CATTTCTACATCTTGAAACCGCCATGACCAACCACTCAGATAGCTAAATCTCCAT 918
390 CACTTACTATTGTTGTTAACTTCACTGACCAACCACTTCAAAATAATCAAACTTGCCAT 449
919 GCCGTTAAACACACCCCGGCGAGTTTGGAGTTTCTTCCCGGAGAGCCGAGACCAATC 978

RESULT 10

US-10-684-651-3

; Sequence 3, Application US/10684651

; Patent No. 6828491

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

Db 450 ACCCGTCAACAACTTGGCAGATATGATGATTTCTTCTTATCTAGCACTCAAGCCCAACA 509
Qy 979 ATCTACTTTCAGGGCTTTCAGCAGGAATACGTTTGGAGGCCCTTCAATCGGAATTCAA 1038
Db 510 GTCCTACTTTCAGAGCTTTCAGCCATAATATTTCTAGAGACCTCTTCCATAGCGAATCGA 569
Qy 1039 TGAGATACGGAGGCTGCTTTAGAGAGATGCAAGGAGGTGAGCAAGAGGAGAGGGCA 1098
Db 570 GGAGATAAACAGGGTTTGTGTTGGAGAGGAAGAGGAGCAGAGCAGCAAGAG----- 621
Qy 1099 GAGCGATGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTAGTAGTCAAGTGTCAAA 1158
Db 622 -----GGAGTGATCTGGAACTCTCAAA 644
Qy 1159 GGAGCACGTTTGAAGAACTTACTAAGCACGCTAAATCCGTCTCAAGAAAGGCTCCGAAGA 1218
Db 645 GGAACAAATTCGGCAACTGAGCAGACGTCGCAAAATCTAGTTCAAGGAAACCACTTTCTCTC 704
Qy 1219 AGAGGAGATATCACCAACCAATCACTTTAGAGAGGAGGAGCCGATCTTTCTTAACAA 1278
Db 705 CGAAGATG-----AACCATTTCAACTTTGAGAAGCCGCAACCCCATCTATTCCAACAA 755
Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAACCCCGAGCTTCAGGACCTGGA 1338
Db 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCAAGCTTCGGGACTTGGGA 812
Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCACACTTCAACTC 1398
Db 813 TATCTCTCTCAGTTCTGTGATATCAACGAAGGAGCTCTTCTTCTACCACTTCAATTC 872
Qy 1399 AAAGGCCATGTTATCGTCTGCTCAACAAAGGAACCTTGAACCTTCGTCGCTGT 1458
Db 873 AAAGGCCATGTTATGATGATTTAATGAAGGAGATGCAAAACATTTGAATTTGTCGAT 932
Qy 1459 AAGAAAGAGCAACACAGAGGCGGCGGGAAGAGGAGGAGGACGACGACGAAGAAGA 1518
Db 933 TAAAGAAACAACAGAGCAGAAACAGGAAGGAACCTTTG----- 975
Qy 1519 GGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAAGCGGATGTT 1578
Db 976 -----GAAGTCAAGGTTACAGAGCTGAAATTTGCTCAAGACGATGATT 1019
Qy 1579 CATCATGCGAGAGCTCATCCAGTAGCCATCAAGCTTCTCCGAATCCCATCTGCTTGG 1638
Db 1020 TGAATTTCCAGCAGCTTATCCATTTGCTGCAACGCTACCTCAACACCTCAATTTCTTGC 1079
Qy 1639 CTTTCGGTATCAACGCTGAAACCAACACAGAACTCTTCTTGCAGGTGATAAGGACAATGT 1698
Db 1080 TTTTGGTATCATGCTGAGAAACCAACAGAGAACTCTTCTTGCAGGCGAAGAACAAATGT 1139
Qy 1699 GATAGACCGATAGAGAACGCAAGGATTTAGCAATTCCTTGGGTGCGGTGAACAAAT 1758
Db 1140 GGTAAAGCAGATAGAAAGCAAGTGCAGGAGCTTGGTTCCTTGGGTCTGCACAAATGT 1199
Qy 1759 TGAGAGCTCATCAAAACCAAGAGAACTCTATTTGTGAGTGTCTGCTCTCA 1812
Db 1200 TGAGAGGCTATTAAAGAGCAGAGGGAATCCTACTTTGTTGATGCTCAGCCTCA 1253

OPERATING SYSTEM: MICROSOFT WINDOWS 95
 SOFTWARE: MICROSOFT WORD 7.0A
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/108,010B
 FILING DATE: 30-Jun-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/019,940
 FILING DATE: JUNE 14,1996
 ATTORNEY/AGENT INFORMATION:
 NAME: LYNNE M. CRISTENBURY
 REGISTRATION NUMBER: 30,971
 REFERENCE/DOCKET NUMBER: BB-1071-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-5481
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1818 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-108-010B-1

Query Match 19.9%; Score 404; DB 3; Length 1818;
 Best Local Similarity 61.2%; Pred. No. 1.7e-112;
 Matches 765; Conservative 0; Mismatches 410; Indels 75; Gaps 4;

Qy	563	AACAAACCTTTCTACTTCCCGTCAAGCGGTTTAGCAGCCGCTAGCGGAAACCAACCGGT	622
Db	577	AAGAAACCTTTCTCTCGGCTCTAACAGGTTGAAACTCTCTTCAAAAACCAATATAGGT	636
Qy	623	AGGATCCGGGCTCTGACAGAGTTTGACCAAAAGGTTCAAGGCTTTCAGAACTCTCCAGAA	682
Db	637	CGCATTCGCGTCTCCAGAGGTTCAACCAACGCTCCCAACAACTTCAGAACTCTCGAGAC	696
Qy	683	CACGATTTGACAGATCGAGCGCAACCTTAACACTCTTTGTTCTTCCCAAGCAGCGTGAT	742
Db	697	TACCGCATTTTGGAGTTCAACTCCAAACCCCAACACCTCTCTCTCCCAACCATGCTGAC	756
Qy	743	GCTGATACATCTTGTATCCAGCAAGGCAAGCCCGTACCGTAGCAATGGCAAT	802
Db	757	GCTGATACCTCATCTGTTATCTTAAACGGAGTCCGCTATCTTCTGTTGTAACCAACGAC	816
Qy	803	AACAGAAAGCTTTAATCTTGAAGGCGCATGCACTCAGAACTCCCATCCGTTTCATT	862
Db	817	GACAGAGACTCTACAGACTTCAATCTGCTGATGCGCTGAGAGTCCCTCAGGAACCA	876
Qy	863	TCCTACATCTTGAACCGCCATGACAAACAGAACTCTGAGATGAGTAAATCTCCATGCC	922
Db	877	TACTATGTTGGTCAACCTTGACAAACCAACCAAAATCTCAGATTAATAACACTCGCATACC	936
Qy	923	GTTAACACACCGCCAGTTTGAAGTTTCTTCCGCGAGCAGCGGAGCAATCATCC	982
Db	937	GTTAACAGCCTGGTAGATTGAGAGTTTCTCTATCTAGCATGAGCTCAACAATCC	996
Qy	983	TACTTGCAGGCTTCAGCAGGAATACGTTGAGGCGGCTTCAATGCGGAATTCAAATGAG	1042
Db	997	TACTTGCAGGATTCAGCAGGAACATTTTGAAGGCTCTTACGATACCAATTCAGGAG	1056
Qy	1043	ATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGAGCAAGAGGAGAGGAGGAGG	1102
Db	1057	ATAAACAGGTTCTGTTTAGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGCTG---	1113
Qy	1103	CGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTGATAGTCAAGTGTCAAGGAG	1162
Db	1114	-----CAAGAGAGCGTGATTTGGAAATCTCGAAGAA	1146
Qy	1163	CACGTTGAAGAACTTACTAAGCAGCTAAATCCGCTCTCAAGAAAGGCTCCGAGAGAGAG	1222
Db	1147	CAGATTCCGGCACTGACGCAACAGTCCCAATCTAGTTCAGAGAAACCAATTTCTTCTGAA	1206

RESULT 12

US-09-758-652-1
 ; Sequence 1, Application US/09758652
 ; Patent No. 6703544
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; APPLICANT: GARY MICHAEL FADER
 ; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 ; CLASSES OF SOYBEAN SEED
 ; PROTEIN GENES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 ; STREET: 1007 MARKET STREET
 ; CITY: WILMINGTON
 ; STATE: DELAWARE
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 19898
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95
 ; SOFTWARE: MICROSOFT WORD 7.0A
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/758,652
 ; FILING DATE: 11-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/019,940
 ; FILING DATE: JUNE 14,1996
 ; ATTORNEY/AGENT INFORMATION:

Db	1315	TTCTCAGTATTGTGATTAACAGGAGGAGCTTCTTCTTACCACTTCAATTCAAAG	1374
Qy	1403	GCATGGTTATCGTCGTCTCAACAAAGGAACCTTGAACCTCGGGCTGTAAGA	1462
Db	1375	GCATAGTGATACCTGGTAAATTAATGAAGGAGATGCAACATTGCACTTGTGGCTAAAA	1434
Qy	1463	AAGACCAACACAGAGGGGCGCGGGAAGAAGAGGAGGACGACGACGAAAGAGGAG	1522
Db	1435	GAACAACACAGAGCAGCAACAGGAAGACCAACCTTTG	1473
Qy	1523	GGAAAGTAACAGAGAGTGGGTAGGTACACAGCGAGGTGAAGAGGCGCATGTGTCATC	1582
Db	1474	-----GAAGTGGGAATATAGAGCCGAATGTCTGAAACAATATTTGTA	1521
Qy	1583	ATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCCGAATCCCATCTGCTTGGCTTC	1642
Db	1522	ATCCACAGCAGGTATCCAGTTGTGGTCAACGCTACCTCAAAATCTGAAATTTCTTTGCTATT	1581
Qy	1643	GGTATCAACGCTGAAACAACCAACAGAACTTCTCTTGCAGGTGATAAGGACAATGTGATA	1702
Db	1582	GGTATTATGCCGAGAACCAACAGAGAACTTCTCGAGGTTCCGACAAATGTGATA	1641
Qy	1703	GACCATAGAGAAGCAAGCAGGATTTAGCAATTCCTGGGTGGGTGAACAAGTTGAG	1762
Db	1642	AGCCAGATACCTAGTCAAGTGCGAGGAGCTTGCAATCCCTGGGTCTGCACAAGCTGTTGAG	1701
Qy	1763	AGCTCATCAAAACACAGAGGAATCTCACTTTGTGAGTGTCTGTCCTCA	1812
Db	1702	AAAGTATTAAAGAACCAAGAGAATCCTACTTTGTGGATGCTCAGCCTAA	1751
RESULT 13			
US-10-684-651-1			
; Sequence 1, Application US/10684651			
; Patent No. 6828491			
; GENERAL INFORMATION:			
; APPLICANT: ANTHONY JOHN KINNEY			
; GARY MICHAEL FADER			
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC			
; CLASSES OF SOYBEAN SEED			
; PROTEIN GENES			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY			
; STREET: 1007 MARKET STREET			
; CITY: WILMINGTON			
; STATE: DELAWARE			
; COUNTRY: UNITED STATES OF AMERICA			
; ZIP: 19898			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: DISKETTE, 3.50 INCH			
; COMPUTER: IBM PC COMPATIBLE			
; OPERATING SYSTEM: MICROSOFT WINDOWS 95			
; SOFTWARE: MICROSOFT WORD 7.0A			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/684,651			
; FILING DATE: 14-Oct-2003			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/09/758,652			
; FILING DATE: 11-Jan-2001			
; APPLICATION NUMBER: 60/019,940			
; FILING DATE: JUNE 14,1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: LYNNE M. CHRISTENBURY			
; REGISTRATION NUMBER: 30,971			
; REFERENCE/DOCKET NUMBER: BB-1071-A			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 302-992-5481			
; TELEFAX: 302-773-0164			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			

```

;
; LENGTH: 1818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID
US-10-694-651-1

```

Query Match 19.9%; Score 404; DB 4; Length 1818;
Best Local Similarity 61.2%; Pred. No. 1.7e-112;
Matches 765; Conservative 0; Mismatches 410; Indels 7

Matches	765,	Conservative	0;	Mismatches	410;	Indels	75;	Gaps	4;
Qy	563	AACAAACCTTTCTACTTCCCGTCAAGCGGTTT	AGCACC	CGCTACGGAAACCAAAACGGT	622				
Db	577	AAGAAACCTTTCTCTTCCGGCTTAAACAGGTT	TCGAAA	CTCTCTTCAAAAACCAATATGGT	636				
Qy	623	AGGATCGGGTCTGACAGAGTTTGACCAAGAGT	CAAGGG	TCAGGAATCTCCAGAAATCTCCAGAAAT	682				
Db	637	CGCAATTCGCGTCTCCAGAGGTTCAACCAACG	CTCCCA	CAACTTCCAGAAATCTCCCGAGAC	696				
Qy	683	CACCGTATTGTGCAGATCGAGGCCAAACCT	TAACACT	CTTGTCTTCCCAAGCAGCGTGTAT	742				
Db	697	TACCGCATTTTGGAGTTCAACTCCAACCCCA	CAACCT	CTCTTCCCCCAACCATGCTGTAC	756				
Qy	743	GCTGATAACATCTCTTGTATTACCAAGGCA	AGGCCA	CCGTCACCGTAGCAAAATGGCAAT	802				
Db	757	GCTGATTACCTCATCGTTATCTTAAACGGAG	CTGCCAT	TTCTTCTCTGGTGAAACAACGAC	816				
Qy	803	AACAGAAAGAGCTTTAATCTTTGACGAGGG	CCATGCA	CTCAGAAATCCCATTCGGTTTCAAT	862				
Db	817	GACAGAGACTCTTACAGACTTCAATCTGGT	GTATCC	CTCGAGAGTCCCTCAGAGAACCAACA	876				
Qy	863	TCCTACATCTTGAACCCGCCATGACCAACAG	AACCTC	AGAGTAGCTAAAACTCCATGCCC	922				
Db	877	TACTATGTGTCAACCTGTACAACACGAAAT	CTCAGAT	TAAACAACCTGCCCATATCCC	936				
Qy	923	GTTTAAACACCCCGCCAGTTTGTAGAGATT	TTCTTCC	GGCGAGCAGCCGAGACCAATCATCC	982				
Db	937	GTTTAAACAGCTTGGTAGATTTGAGAGTTT	CTTCTCT	ATCTAGCACTGAAGCTCAACAATCC	996				
Qy	983	TACTTGCAGGGCTTCACGAGGAATACGTT	TGGAGC	CCGCTTCAATTCGGGAATTCGAATGAG	1042				
Db	997	TACTTGCAGAGATTACGACAGGAACATTT	TAGAGC	CTCTACGATACCAAAATTCGAGGAG	1056				
Qy	1043	ATACGAGGGTGTCTTTAGAAGAAATGCA	GAGGAT	GTGACGAAGGAGAGAGGGCAGAGG	1102				
Db	1057	ATAAACAGGTTCTGTTTGTAGAGAGAG	GGGCGA	CGACAGGGGAGCAGAGGCTG----	1113				
Qy	1103	CGATGGAGTACTCGGAGTAGTGAAACAAT	GAAGAG	TGATGATGCAAGGTGTCAAAGGAG	1162				
Db	1114	-----	-----	CAAGAGAGCGTGATTGTGGAAATCTCGAAGGAA	1146				
Qy	1163	CAGTTTGAAGAACTTACTAAGCAGCTAAAT	CCGTAA	ATCTCAAAGAAGGCTCCGAAGAAGAG	1222				
Db	1147	CAGATTCCGGCACTGACGCAAAACGTC	CAAAAT	CTAGTTTCAAGGAAAAACATTTCTTCTGAA	1206				
Qy	1223	GGAGATATCACCACCCCAATCAACTTC	GAGAGA	AGCGGACCGCATCTTTCTTAAACAATTT	1282				
Db	1207	G-----	-----	ATAAACCTTTTAACTTGAGAGCCGCAACCCCATCTACTCCACACAGCTT	1257				
Qy	1283	GGGAAGTTATTTGAGGTGAAGCCACAGACA	AGAAAC	CCCCCAGCTTCAGGACCTCGGACATG	1342				
Db	1258	GGCAAGTTCTTTGAGATACACCCCAAG	-----	AGAAAAACCCCAAGCTTCGGGACTTCGGATATC	1314				
Qy	1343	ATGCTACCTGTGTAGAGATCAAGAAGGAG	CTTTGAT	GTCTCCCACTTCAACTTCAAG	1402				
Db	1315	TTCCCTCAGTATTGTGGATATGAACGAG	GGGAGC	TTCTTCTTACACACTTCAATTTCAAAG	1374				
Qy	1403	GCCATGTTATCGTCGTCGTCAACAAAGGA	ACTCGA	AAACCTTGAACCTCGTGCGCTGAAGA	1462				
Db	1375	GCGATAGTGATCTGGTAAATTAAATGA	AGGAGAT	CGAAACATTTGAACTTTGTTGGCCTTAAA	1434				
Qy	1463	AAAGAGCAACAAACAGAGGGGACGGCG	GGGAAAG	AGAGGAGGACGAGAACGAGAGAGGAG	1522				

Query Match 16.0%; Score 324.2; DB 3; Length 1920;
Best Local Similarity 58.5%; Pred. No. 4.7e-88;

Matches	764;	Conservative	0;	Mismatches	408;	Indels	135;	Gaps	6;
Qy	563	AAACAACCTTTTCTACTTTCCGTC	CAAGCGGTTT	TAGCACCGCTAC	GGGAACCAAAACGGT	622			
Db	625	AAGAACCTTTTCTACTTTCAACTCTAAAGGTTCCAAACTCTCTTC	AAAAACCAATATGGC	684					
Qy	623	AGGATCCGGTCTCTGCAGAGGTTTGA	CAAAAGGTTCAAGGAGTTT	CAGATCTCCAGAAT	682				
Db	685	CACGTTCCGCTCTCCAGAGGTTCAACA	CAAAACGCTCCCAACAGCTTC	CAGATCTCCGAGAC	744				
Qy	683	CACCGTATTGTGCAGATCGAGGCCAAACCTTAA	CACTCTTGTCTTCCAGACGAGCTGAT	742					
Db	745	TACCGCATTTTGGAGTTTCAACTCCAAACCCCAACACCCCTTCTCTCC	CCCACTGCTGAC	804					
Qy	743	GCTGATAACATCTCTGTTTATTCACGAAGGCGAAGCCACCGT	GACGCTAGCAGT	802					
Db	805	GCTGATTTACCTCATCGTTATCTTAA	CGGAGACTGCCATCTTACCTTGGTGAACAACGAC	864					
Qy	803	AACAGAAAGAGCTTTAATCTTTGACGAGGGCCATG	CACTCAGAAATCCC-	849					
Db	865	GACCGAGACTCTTTACAACACCTTCAATCTGGCGATG	CCCTTAAGAGTCCCTGCAGGAACCAACA	924					
Qy	850	-	-	-	-	-	-	-	-
Db	925	TTCTATGTGTTTAA	CCCTGCAACACGACGAGAAATCTCAGAAATGATAGCAGGAACCAACATTC	984					
Qy	866	TACATCTTGAACCGCATGATCAACACCAACCTCAGAGT	AGTCTAAAATCTCCATGSCCGGTT	925					
Db	985	TATGTGTTTAACTCTGACACGACGAGAAATCTCAGAAATGATA	CACATCGGCATACCCGTT	1044					
Qy	926	AACACACCCCGCCAGTTTGTAGGATTTCTTCCCGG	CAGACCGCAGACCAATCATCTCATC	985					
Db	1045	AACAAACCCCGTAGATTGTGAGGTTTCTTCTATCTAGCACTCAAGCTCAACAGTCTCATC	1104						
Qy	986	TTGCAGGGCTTACGACGAGATACGTTGAGAGCGCCCTTCAATCGG	AAATTCATAGATA	1045					
Db	1105	TTGCAAGGTTTCAGCAAGAAATATTTCTAGAGCCCTCATACGACACCAAAATTCGAGGAGATA	1164						
Qy	1046	CGAGGGTGTGTTTAGAAGAGAAATGCAAGGAGGTGAGCAAGAGAGAGAGGCGCAGAGSCGA	1105						
Db	1165	AACAGGTTCTGTTTGTGTAGAGAGGGGCGACGCAACAGGGGAGGAGCGCTGC	-	1219					
Qy	1106	TGGAGTACTCGGAGTGTGTGAAACAATGAAGGAGTGTATGTCAAGTGTCAAAAGGAGCAC	1165						
Db	1220	-	-	-	-	-	-	-	-
Qy	1166	GTTGAAGAACTTACTAAGCAGCTTAAATTCGTTCAAAAGAAAGGCTCCGAAAGAGGGGA	1225						
Db	1255	ATTCGGGAACTGAGCAAAACATGCCAAATCTAGTTCAAGGAAACCAATTTCTTCTGAAG--	1312						
Qy	1226	GATATCACCACCAATCAACTTTGAGAGAGGGGAGGCCCATCTTTTCTACAACTTTGGG	1285						
Db	1313	-	-	-	-	-	-	-	-
Qy	1286	AAGTTATTCGAGGTGAAGCCAGACAAGAAAGAACCCCGCAGCTTCAGGACCTGGACATGATG	1345						
Db	1366	AGTTGTTTGATTTACCC-	-	-	-	-	-	-	-
Qy	1346	CTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGTCCACACTTCAACTCAAAGGCC	1405						
Db	1420	CTCAGTGTGTGGATAAGAACGAGGGAGCTCTTTTCTACACACTTCAATTCAAAGGCC	1479						
Qy	1406	ATGGTTATCGTCGTCTCAACAAAGGAACTGGAAACCTTGAATCTCGTGGCTGTGAAGAAA	1465						
Db	1480	ATAGTGTGTAAGTAGTGAATTAAGAGGAGAGCAAAACATTGAACTTGT--	-	-	-	-	-	-	-
Qy	1466	GAGCAACAAACAGAGGGGACCGCGGGAAGAGAGGAGGACGAAGACGAAGAGGAGGGGA	1525						
Db	1537	GAACAAACACAGGCGACGACAAAGGAAGCAACCTTTGGA	-	-	-	-	-	-	-
Qy	1526	AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAAGCGATGTGTTCAATCATG	1585						
Db	1579	-	-	-	-	-	-	-	-

Qy	1586	CCAGCAGCTCATCCAGTAGGCCATCAAGCTTCTCCGAACTCCATCTGCTTGGCTTCGGT	1645
Db	1627	CCAGCAGGTTATCCAGGTTATGGTCAACGCCTCAGATCTGAATTTCTTTGCTTTTGGT	1686
Qy	1646	ATCAAGCCTCAAAACACCAACAGAACTCTTCCTTGCAGGTCATAAGGACAATGTGATAGAC	1705
Db	1687	ATCAATGCCAGAACCAACGAGGAACCTTCCTTGCAGGTTCGAAAGACAATGTGATAAGC	1746
Qy	1706	CAGATAGAGAAGCAAGCGAAGGATTTAGCAATTCCTCGGTCGGGTGAAACAAGTTGAGAAG	1765
Db	1747	CAGATCCTAGTCAAGTCGAGGAGCTTGCCTTCCCTAGGTCCTCAAAAGATATTGAGAAC	1806
Qy	1766	CTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGCCTCGTCTCTCA	1812
Db	1807	CTAATAAGAGCAAAAGTGAGTCTACTTTGTGGATGCTCAGGCTCA	1853
RESULT 15			
US-09-758-652-2			
; Sequence 2, Application US/09758652			
; Patent No. 6703544			
; GENERAL INFORMATION:			
; APPLICANT: ANTHONY JOHN KINNEY			
; GARY MICHAEL FADER			
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC			
; CLASSES OF SOYBEAN SEED			
; PROTEIN GENES			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY			
; STREET: 1007 MARKET STREET			
; CITY: WILMINGTON			
; STATE: DELAWARE			
; COUNTRY: UNITED STATES OF AMERICA			
; ZIP: 19898			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: DISKETTE, 3.50 INCH			
; COMPUTER: IBM PC COMPATIBLE			
; OPERATING SYSTEM: MICROSOFT WINDOWS 95			
; SOFTWARE: MICROSOFT WORD 7.0A			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/758,652			
; FILING DATE: 11-Jan-2001			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 60/019,940			
; FILING DATE: JUNE 14,1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: LYNNE M. CHRISTENBURY			
; REGISTRATION NUMBER: 30,971			
; REFERENCE/DOCKET NUMBER: BB-1071-A			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 302-992-5481			
; TELEFAX: 302-773-0164			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1920 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: cdna			
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:			
US-09-758-652-2			

	Query Match	16.08;	Score 324.2;	DB 4;	Length 1920;
	Best Local Similarity	58.5%;	Pred. No. 4.7e-88;		
	Matches 764;	Conservative	0;	Mismatches 408;	Indels 135; Gaps 6;
Qy	563	AAACAACCTTTCTACTTCCCGTCAAGCGCGTTTAGCACCCGCTACGGGAACCAAAACGGT	622		
Db	625	AAGAAGCCCTTTTCACTTCAACTTAAAGGTTCCAAACTCTTTCAAAAACCAATATGGC	684		

```
Qy 623 AGGATCCGGTCTTGACAGAGTTTGACCAAGGTCAAGGAGTTTCAGAAATCCAGAAT 682
Db 685 CACGTTCCGTCCTCCAGAGTTTCAAAAGCTCCCAACAGCTTCAGAAATCTCCGAGAC 744
Qy 683 CACCGTATGTGAGATCAGAGCCAAACTTAACACTCTTGTCTTCCCAAGCAGCTGAT 742
Db 745 TACCGCATTTTGGAGTTCAACTCCAAACCCCAACACCTTCTTCCGCCCAACCATGCTGAC 804
Qy 743 GCTGATACATCTTGTATCAGCAAGGGCAAGCCAGCTGACCGTAGCAAAATGGCAAT 802
Db 805 GCTGATTAACCTCATCTGTTATCTTAACGGGACTGCCATCTTACCTTTGGTGAACAACGAC 864
Qy 803 AACAGAAAGACTTTAATCTTGAAGGGCCATCACTCAGAATCCC----- 849
Db 865 GACCGAGACTCTTACAACCTTCAATCTGGCGATGCCCTAAGAGTCCCTGCAAGGAACACA 924
Qy 850 -----ATCCGGTTTCATTTCC 865
Db 925 TTCTATGTGTTAAACCTTGACAACGAGAGAACTCTCAGAAATGATAGCAGGAACCAATTC 984
Qy 866 TACATCTTTGAACCGCATGACAACAGAACTCAGAGTAGTAAATCTCATGSCCGTT 925
Db 985 TATGTGTTAAACCTTGACAACGAGAGAACTCTCAGAAATGATAACACTCGCCATACCCGTT 1044
Qy 926 AACACACCCGCCAGTTTGGAGTTTCTTCCCGGAGCAGCCGAGACCAATCATCCTAC 985
Db 1045 AACAAACCCGTAGATTGAGAGTTCTTCTATAGCACTCAAGCTCAACAGTCTCCTAC 1104
Qy 986 TTGAGGGCTTCAGCAGGAATACGTTGGAGGCCCTTCAATGCGGAATTCATAGATA 1045
Db 1105 TTGCAAGGGTTTCAGCAAGAATATTCTAGAGGCTCATACGACACCAATTCGAGAGATA 1164
Qy 1046 CGGAGGTGCTGTTAGAAGAAATGACAGAGGTGAGCAAGAGAGAGAGGGCAGAGCGA 1105
Db 1165 AACAAAGTTCTGTTTGGTAGAGAGGGGCGAGCAACAAGGGGAGAGAGGCTGC----- 1219
Qy 1106 TGGAGTACTCGGAGTAGTGAGAAACAATGAAGAGGTAGTCAAAAGTGTCAAGGAGCAC 1165
Db 1220 -----NAGAGAGTGTGATTTGGAAATCTCAAGAAACAA 1254
Qy 1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAGAAAGGCTCCGAAGAAAGGGA 1225
Db 1255 ATTCGGGAACCTGAGCAAACTGCCAAATCTAGTTCAAGGAAACCAATTTCTCTGAAG-- 1312
Qy 1226 GATATCAACCAACCAATCACTTGAGAGAGGGCAGCCGATCTTTCTAACAACTTTGGG 1285
Db 1313 -----ATAAACCTTTCAACTTGGGAAGCGCGACCCCATCTATTCCAAACAGCTTGGC 1365
Qy 1286 AAGTTATTGAGGTGAAGCCAGACAAGAAAGAACCCAGCTTCAGGACCTGGACATGATG 1345
Db 1366 AAGTTGTTGAGATTACCC-----AGAGAAACCTCAGCTTCGGGACTTGGATGTCTTC 1419
Qy 1346 CTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCC 1405
Db 1420 CTCAGTTGTGGGATATCAACGAGGGAGCTCTTTTCTACCACTTCAATTCAGAGGCC 1479
Qy 1406 ATGTTATGTCGTGTCACAAAGGAACCTGGAACCTTGAACTCGTGGCTGTGAAGAAA 1465
Db 1480 ATAGTGTACTAGTGTATTAATGAAGGAGAACCAACATTTGAATTTGT---TGGCATTAAA 1536
Qy 1466 GAGCAACACAGAGGGGCGCGGAGAGAGAGGAGGAGCAGAGCAAGAGAGAGAGGGA 1525
Db 1537 GAACAAACACAGAGGAGCAGCAAGAGAGCAACCTTTGGAA----- 1578
Qy 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAGGCGATGTTCATCATG 1585
Db 1579 -----GTGCGGAATATAGAGCTGAATTTGTCTGAACAAGATATATTGTAATC 1626
Qy 1586 CCAGCAGTCAATCCAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db 1627 CCAGCAGGTTATCCAGTTATGGTCAACGCTACCTCAGATCTGAATTTCTTTGCTTTTGGT 1686
Qy 1646 ATCAACGCTGAACCAACCAACAGAAATCTTCTTTCAGGTGATAGGACAATGTGATAGAC 1705
```

```
Db 1687 ATCAATGCCGAGAACCAACGAGGAACCTTCTTGACAGTTTCGAAAGACAATGTGATAAGC 1746
Qy 1706 CAGATAGAGAAAGCAAGCGAAGGATTTTAGCAATTCCTGGGTGCGGTGAACAAGTTGAGAAG 1765
Db 1747 CAGATACCTAGTCAAGTGCAGGAGCTTGCCTTCCCTAGGTCTGCAAAAGATATTGAGAAC 1806
Qy 1766 CTCAATCAAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
Db 1807 CTAATAAGAGCCAAAGTGAGTCTCTACTTTGTGGATGCTCAGCCTCA 1853
```

Search completed: August 24, 2005, 03:56:27
Job time : 348.307 secs

This Page Blank (uspiu)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 1228.33 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataatcatatatttcac.....cgttgtgcgtgtttctcc 2032

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2032	100.0	2032	10	US-09-731-375A-1
2	2032	100.0	2032	14	US-10-728-806-1
3	2032	100.0	2032	17	US-10-100-303A-5
4	2032	100.0	2032	20	US-10-728-323-1
5	2032	100.0	2032	20	US-10-728-051-1
6	2032	100.0	2032	21	US-10-899-551-1
7	2032	100.0	2032	21	US-10-958-324-5

8	1752.4	86.2	1949	17	US-10-100-303A-6	Sequence 6, Appli
9	406.8	20.0	1254	17	US-10-245-227B-15	Sequence 15, Appli
10	406.8	20.0	1278	17	US-10-245-227B-13	Sequence 13, Appli
11	406.8	20.0	1320	9	US-09-758-652-3	Sequence 3, Appli
12	406.8	20.0	1320	17	US-10-245-227B-3	Sequence 3, Appli
13	406.8	20.0	1320	18	US-10-684-651-3	Sequence 3, Appli
14	406.8	20.0	1320	19	US-10-757-074-3	Sequence 3, Appli
15	406.8	20.0	1320	19	US-10-757-155-3	Sequence 3, Appli
16	406.8	20.0	1320	19	US-10-757-667-3	Sequence 3, Appli
17	404	19.9	1818	9	US-09-758-652-1	Sequence 1, Appli
18	404	19.9	1818	18	US-10-684-651-1	Sequence 1, Appli
19	404	19.9	1818	18	US-10-757-074-1	Sequence 1, Appli
20	404	19.9	1818	19	US-10-757-155-1	Sequence 1, Appli
21	404	19.9	1818	19	US-10-757-667-1	Sequence 1, Appli
22	404	19.9	2073	20	US-10-739-930-3296	Sequence 3296, Ap
23	404	19.9	2260	18	US-10-424-599-10353	Sequence 10353, A
24	404	19.9	5642	18	US-10-424-599-10364	Sequence 10364, A
25	391	19.2	5642	18	US-10-424-599-10364	Sequence 10364, A
26	390	19.2	1656	18	US-10-425-114-9571	Sequence 9571, Ap
27	374.2	18.4	2031	18	US-10-424-599-117263	Sequence 117263,
28	324.2	16.0	1920	9	US-09-758-652-2	Sequence 2, Appli
29	324.2	16.0	1920	18	US-10-684-651-2	Sequence 2, Appli
30	324.2	16.0	1920	19	US-10-757-074-2	Sequence 2, Appli
31	324.2	16.0	1920	19	US-10-757-155-2	Sequence 2, Appli
32	324.2	16.0	1920	19	US-10-757-667-2	Sequence 2, Appli
33	293.4	14.4	1107	18	US-10-425-114-8117	Sequence 8117, Ap
34	279.8	13.8	1088	18	US-10-425-114-7645	Sequence 7645, Ap
35	279.4	13.7	1083	18	US-10-425-114-14774	Sequence 14774, A
36	279.2	13.7	1079	18	US-10-425-114-8143	Sequence 8143, Ap
37	267	13.1	1037	18	US-10-425-114-14858	Sequence 14858, A
38	263.8	13.0	1034	18	US-10-425-114-14773	Sequence 14773, A
39	260.4	12.8	1024	18	US-10-425-114-7686	Sequence 7686, Ap
40	247.8	12.2	1011	18	US-10-425-114-8555	Sequence 8555, Ap
41	233	11.5	951	18	US-10-425-114-8579	Sequence 8579, Ap
42	214.8	10.6	863	18	US-10-425-114-7590	Sequence 7590, Ap
43	214.8	10.6	880	18	US-10-425-114-7667	Sequence 7667, Ap
44	214.8	10.6	934	18	US-10-425-114-6929	Sequence 6929, Ap
45	199.6	9.8	720	18	US-10-425-114-15167	Sequence 15167, A

ALIGNMENTS

RESULT 1
US-09-731-375A-1
; Sequence 1, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731.375A
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-09-731-375A-1

Query Match	100.0%	Score 2032;	DB 10;	Length 2032;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2032;	Conservative	0;	Mismatches	0; Indels
				0; Gaps
				0;
Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Db	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60	
Qy	61	GGTTCTCCACTGCTGTGCTAGGATCCTTGCTGCTTCAGTTCTGCAACGCA	120	

Db 61 GTTTCCTCACTGATGCTGTGTCTAGGGATCCTTGTCTCGCTTCAGTTCCTCGCAACGCA 120
Qy 121 TGCCAAAGTCACTCACTTACAGAGAAAGAAACAGAGAACCCCTGCGCCAGAGGTCCTCCCA 180
Db 121 TGCCAAAGTCACTCACTTACAGAGAAAGAAACAGAGAACCCCTGCGCCAGAGGTCCTCCCA 180
Qy 181 GAGTGTGTCAACAGGAACCGGATGATCTTGAAGCAAAAGGCGATGCGAGTCTCGCTGCACCA 240
Db 181 GAGTGTGTCAACAGGAACCGGATGATCTTGAAGCAAAAGGCGATGCGAGTCTCGCTGCACCA 240
Qy 241 GCTCGAGTATGATCCTCGTGTGTCTATGATCTCTGAGGACACATGCGACACCAACCA 300
Db 241 GCTCGAGTATGATCCTCGTGTGTCTATGATCTCTGAGGACACATGCGACACCAACCA 300
Qy 301 ACGTTCCTCCAGGGGAGCGGACACGTGCGCGCCACCCGAGACTACCGATGATGACCG 360
Db 301 ACGTTCCTCCAGGGGAGCGGACACGTGCGCGCCACCCGAGACTACCGATGATGACCG 360
Qy 361 CCCTCAACCCCGAAGAGGAGGAGCGGCGATGCGGACCAAGCTGGACCCGAGGGAGCGTGA 420
Db 361 CCCTCAACCCCGAAGAGGAGGAGCGGCGATGCGGACCAAGCTGGACCCGAGGGAGCGTGA 420
Qy 421 AAGAGAAAGACTGAGAGACCAACAGAGAAAGATTTGGAGCGGACCAAGTCAATCAGCAGCC 480
Db 421 AAGAGAAAGACTGAGAGACCAACAGAGAAAGATTTGGAGCGGACCAAGTCAATCAGCAGCC 480
Qy 481 ACGGAAATTAAGGCCCGAAGAGAGAGAGAAAGAGTGGGGAACACACAGGTAGCCCA 540
Db 481 ACGGAAATTAAGGCCCGAAGAGAGAGAGAAAGAGTGGGGAACACACAGGTAGCCCA 540
Qy 541 TGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Qy 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCGAGAGTTTGACCAAGAGTCAAG 660
Db 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCGAGAGTTTGACCAAGAGTCAAG 660
Qy 661 GCAGTTTCAGAACTCTCAGAAATCACCGTATTTGTGCAGATCGAGGCCAAACCTAACACTCT 720
Db 661 GCAGTTTCAGAACTCTCAGAAATCACCGTATTTGTGCAGATCGAGGCCAAACCTAACACTCT 720
Qy 721 TGTTCCTTCCCAAGCACGCTGATGCTGATAACATCTCTTGTATCCAGCAAGGGCAAGCCAC 780
Db 721 TGTTCCTTCCCAAGCACGCTGATGCTGATAACATCTCTTGTATCCAGCAAGGGCAAGCCAC 780
Qy 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAACTTTGACGAGGGCCATGCACT 840
Db 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAACTTTGACGAGGGCCATGCACT 840
Qy 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGGCATGACAAACAGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGGCATGACAAACAGAACTCAG 900
Qy 901 AGTAGCTAAATCTCCATGCGGTAAACACACCGCGGAGTTTGAAGATTTCTTCCCGGC 960
Db 901 AGTAGCTAAATCTCCATGCGGTAAACACACCGCGGAGTTTGAAGATTTCTTCCCGGC 960
Qy 961 GAGCAGCGGAGACCAATCTACTTTCAGAGGCTTCAGCAGGAATACGTTGAGGCGCG 1020
Db 961 GAGCAGCGGAGACCAATCTACTTTCAGAGGCTTCAGCAGGAATACGTTGAGGCGCG 1020
Qy 1021 CTTCAATCCGGAATTCATGAGATACGAGGGTCTGTTAGAAGAGAAATCAGAGGTGA 1080
Db 1021 CTTCAATCCGGAATTCATGAGATACGAGGGTCTGTTAGAAGAGAAATCAGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGGCAGAGGGGATCGAGTACTCGAGGTAGTGAGAAACAAATGAAGAGT 1140
Db 1081 GCAAGAGGAGAGGGCAGAGGGGATCGAGTACTCGAGGTAGTGAGAAACAAATGAAGAGT 1140
Qy 1141 GATAGTCAAAAGTGTCAAAAGGAGCAGGTTGAAGACTTACTTAAGCAGCTAAATCCGCTC 1200
Db 1141 GATAGTCAAAAGTGTCAAAAGGAGCAGGTTGAAGACTTACTTAAGCAGCTAAATCCGCTC 1200

Qy 1201 AAAGAAAGGCTCCGAAGAGAGGGAGATATCACAAACCAATCAACTTGAGAGAAGCGCA 1260
Db 1201 AAAGAAAGGCTCCGAAGAGAGGGAGATATCACAAACCAATCAACTTGAGAGAAGCGCA 1260
Qy 1261 GCCCGATCTTTCTAAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAGACCC 1320
Db 1261 GCCCGATCTTTCTAAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAGACCC 1320
Qy 1321 CGAGCTTCAGGACCTCGACATGATCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGAT 1380
Db 1321 CGAGCTTCAGGACCTCGACATGATCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGAT 1380
Qy 1381 GCTCCACACACTTCAACTCAAAAGGCCATGGTTATCGTCGTCAACAAAGGAACCTGAAA 1440
Db 1381 GCTCCACACACTTCAACTCAAAAGGCCATGGTTATCGTCGTCAACAAAGGAACCTGAAA 1440
Qy 1441 CTTTGAATCTGTGGCTGTGAAGAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
Db 1441 CTTTGAATCTGTGGCTGTGAAGAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
Qy 1501 GGACGAAGACGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACACGAGGTT 1560
Db 1501 GGACGAAGACGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACACGAGGTT 1560
Qy 1561 GAAGGAAGGCGATGTCTTCAATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Db 1561 GAAGGAAGGCGATGTCTTCAATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Qy 1621 CGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAACAAACACAGAAATCTTCCTGC 1680
Db 1621 CGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAACAAACACAGAAATCTTCCTGC 1680
Qy 1681 AGGTGATAAGGACAACTGTGATAGCAGATAGAGAGCAAGCAAGGATTTAGCATTCCTC 1740
Db 1681 AGGTGATAAGGACAACTGTGATAGCAGATAGAGAGCAAGCAAGGATTTAGCATTCCTC 1740
Qy 1741 TGCGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGAGAAATCTCAGTTGTGAG 1800
Db 1741 TGCGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGAGAAATCTCAGTTGTGAG 1800
Qy 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCCTGAGAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCACTCCTTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGTCCACTCCTTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTCAGAAATGAGGCAACTTCTTATGATCGATAATAGATCAGCTTTTGTACT 1980
Db 1921 TTTTAACTCAGAAATGAGGCAACTTCTTATGATCGATAATAGATCAGCTTTTGTACT 1980
Qy 1981 CTACTATCAAAAACTTATCAATAAATAAAAGCTTTGTGCGTTGTCTCC 2032
Db 1981 CTACTATCAAAAACTTATCAATAAATAAAAGCTTTGTGCGTTGTCTCC 2032

RESULT 2
US-10-228-806-1
; Sequence 1, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2032

TYPE: DNA		Db		1020	
: ORGANISM: Arachis hypogaea		Qy		1080	
US-10-228-806-1		Db		1080	
Query Match 100.0%; Score 2032; DB 14; Length 2032;		Qy		1140	
Best Local Similarity 100.0%; Pred. No. 0;		Db		1200	
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Qy		1260	
1 AATATCATATATATTCATCATCTATATATAGTAGTAGCAGAGCAATGAGGGAG 60		Db		1320	
1 AATATCATATATATTCATCATCTATATATAGTAGTAGCAGAGCAATGAGGGAG 60		Qy		1380	
61 GGTTCCTCCACTGATGCTGTCTAGGATCCTCTCTGCTTCTGCTTCTGCAACGCA 120		Db		1440	
61 GGTTCCTCCACTGATGCTGTCTAGGATCCTCTCTGCTTCTGCTTCTGCAACGCA 120		Qy		1500	
121 TGCCAAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCACAGAGTGCCTCA 180		Db		1560	
121 TGCCAAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCACAGAGTGCCTCA 180		Qy		1620	
181 GAGTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGTGCACAA 240		Db		1680	
181 GAGTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGTGCACAA 240		Qy		1740	
241 GCTGAGTATGATCCTCTGCTGTCTATGATCTCGAGGACACACTGGCACACCAACCA 300		Db		1800	
241 GCTGAGTATGATCCTCTGCTGTCTATGATCTCGAGGACACACTGGCACACCAACCA 300		Qy		1860	
301 ACGTTCCTCCAGGGAGCGGACACGCTGGCGCCCAACCCGAGACTACGATGATGACCG 360		Db		1920	
301 ACGTTCCTCCAGGGAGCGGACACGCTGGCGCCCAACCCGAGACTACGATGATGACCG 360		Qy		1980	
361 CCGTCAACCCGAGAGAGAGGAGCGGATGGGACACGCTGGAACCGAGGAGCGTGA 420		Db		2040	
361 CCGTCAACCCGAGAGAGAGGAGCGGATGGGACACGCTGGAACCGAGGAGCGTGA 420		Qy		2100	
421 AAGAGAGAGACTGAGAGCAACCAAGAGAGATTGGAGGCGACCAAGTCAATCAGCAGCC 480		Db		2160	
421 AAGAGAGAGACTGAGAGCAACCAAGAGAGATTGGAGGCGACCAAGTCAATCAGCAGCC 480		Qy		2220	
481 ACGGAAATAAGGCCCGAGAGAGAGAGAGCAAGAGTGGGAAACACAGGTAGCCA 540		Db		2280	
481 ACGGAAATAAGGCCCGAGAGAGAGAGAGCAAGAGTGGGAAACACAGGTAGCCA 540		Qy		2340	
541 TGTGAGGAGAGAAACATCTCGGAAACACCCCTTTCTACTTCCCGTCAAGGCGGTTAGCAC 600		Db		2400	
541 TGTGAGGAGAGAAACATCTCGGAAACACCCCTTTCTACTTCCCGTCAAGGCGGTTAGCAC 600		Qy		2460	
601 CCGCTACCGGAAACCAAAACGGTAGGATCCGGGCTCTGAGAGGTTTGACCAAGGTCAG 660		Db		2520	
601 CCGCTACCGGAAACCAAAACGGTAGGATCCGGGCTCTGAGAGGTTTGACCAAGGTCAG 660		Qy		2580	
661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAACCTAACCTCT 720		Db		2640	
661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAACCTAACCTCT 720		Qy		2700	
721 TGTTCCTCCAGAGCAGCTGATGCTGATAACATCTCTGTTATTCAGCAAGGCGCAAGCCAC 780		Db		2760	
721 TGTTCCTCCAGAGCAGCTGATGCTGATAACATCTCTGTTATTCAGCAAGGCGCAAGCCAC 780		Qy		2820	
781 CGTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840		Db		2880	
781 CGTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACT 840		Qy		2940	
841 CAGAAATCCATCCGTTTCATTTCTTCTACATCTTGAACCGCATGACACCAAGACCTCAG 900		Db		3000	
841 CAGAAATCCATCCGTTTCATTTCTTCTACATCTTGAACCGCATGACACCAAGACCTCAG 900		Qy		3060	
901 AGTAGCTAAATCTCCATGCGGTTTAAACACCCCGGCGAGTTGAGGATTTCTTCCCGGC 960		Db		3120	
901 AGTAGCTAAATCTCCATGCGGTTTAAACACCCCGGCGAGTTGAGGATTTCTTCCCGGC 960		Qy		3180	
961 GAGCAGCGGAGCAATCATCTTCTGAGGCTTCAGCAGGAATAGTTGGAGGCGCG 1020		Db		3240	
961 GAGCAGCGGAGCAATCATCTTCTGAGGCTTCAGCAGGAATAGTTGGAGGCGCG 1020		Qy		3300	

US-10-100-303A-5									
; Sequence 5, Application US/10100303A									
; Publication No. US20030202980A1									
; GENERAL INFORMATION:									
; APPLICANT: Caplan, et al.									
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction									
; TITLE OF INVENTION: to Allergy									
; FILE REFERENCE: 2002834-0166									
; CURRENT APPLICATION NUMBER: US/10/100,303A									
; CURRENT FILING DATE: 2002-03-18									
; NUMBER OF SEQ ID NOS: 138									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 5									
; LENGTH: 2032									
; TYPE: DNA									
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1									
US-10-100-303A-5									
Query Match 100.0%; Score 2032; DB 17; Length 2032;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGGAGCAATGAGAGGAG	60						
Db	1	AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGGAGCAATGAGAGGAG	60						
Qy	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120						
Db	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120						
Qy	121	TGCCAAGTCATCACTTACAGAAAGAAAACAGAGAACCCCTCGCCCCAGAGGTGCCTCCA	180						
Db	121	TGCCAAGTCATCACTTACAGAAAGAAAACAGAGAACCCCTCGCCCCAGAGGTGCCTCCA	180						
Qy	181	GAGTTGTCAACAGAAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240						
Db	181	GAGTTGTCAACAGAAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240						
Qy	241	GCTCGAGTATGATCCTCGTTGTGTATGATCTCTCGAGGACACACTGGCACCAACCA	300						
Db	241	GCTCGAGTATGATCCTCGTTGTGTATGATCTCTCGAGGACACACTGGCACCAACCA	300						
Qy	301	ACGTTTCCCTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATGACCG	360						
Db	301	ACGTTTCCCTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATGACCG	360						
Qy	361	CCGTCACCCCGAAGAGAGAGGAGGCGGATGGGACCAAGCTTGAACCGAGGGAGCGTGA	420						
Db	361	CCGTCACCCCGAAGAGAGAGGAGGCGGATGGGACCAAGCTTGAACCGAGGGAGCGTGA	420						
Qy	421	AAGAGAGAGACTGGAGACCAACAGAGAGATTGGAGGCGCAAGTCAACAGCAGCC	480						
Db	421	AAGAGAGAGACTGGAGACCAACAGAGAGATTGGAGGCGCAAGTCAACAGCAGCC	480						
Qy	481	ACGGAATAATAGGCCCGGAG	540						
Db	481	ACGGAATAATAGGCCCGGAG	540						
Qy	541	TGTGAGGAGAAACATCTCGGAACAAACCCCTTTCTACTTCCCGTCAAGGGCGTTTAGCAC	600						
Db	541	TGTGAGGAGAAACATCTCGGAACAAACCCCTTTCTACTTCCCGTCAAGGGCGTTTAGCAC	600						
Qy	601	CCGCTACGGGAACCAAAACGCTAGGATCCGGTCTCGAGAGTTTCAGCAAAAGGTCAAG	660						
Db	601	CCGCTACGGGAACCAAAACGCTAGGATCCGGTCTCGAGAGTTTCAGCAAAAGGTCAAG	660						
Qy	661	GCAGTTTCAGAAATCTCCAGATACCGGTATTTGTGCAGATCGAGGCCAAACCTAACACTCT	720						
Db	661	GCAGTTTCAGAAATCTCCAGATACCGGTATTTGTGCAGATCGAGGCCAAACCTAACACTCT	720						
Qy	721	TGTTCTTCCCAGCAACGCTGATGATTAACATCTCTGTTTATCCAGCAAGGGCAAGCCAC	780						
Db	721	TGTTCTTCCCAGCAACGCTGATGATTAACATCTCTGTTTATCCAGCAAGGGCAAGCCAC	780						

Qy	781	CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCATGCACCT	840
Db	781	CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCATGCACCT	840
Qy	841	CAGAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCCGCCATGACAAACAGAACCTCAG	900
Db	841	CAGAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCCGCCATGACAAACAGAACCTCAG	900
Qy	901	AGTAGCTAAAAATCTCCATGCCCGTTAAACACACCCCGGCCAGTTTGAGGATTTCTTCCGGC	960
Db	901	AGTAGCTAAAAATCTCCATGCCCGTTAAACACACCCCGGCCAGTTTGAGGATTTCTTCCGGC	960
Qy	961	GAGCAGCCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGAGGAGCCGC	1020
Db	961	GAGCAGCCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGAGGAGCCGC	1020
Qy	1021	CTTCAATGCGGAATTCATAGATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA	1080
Db	1021	CTTCAATGCGGAATTCATAGATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA	1080
Qy	1081	GCAAGAGGAGAGGGCAGAGCGATGAGTACTCGAGTAGTGAGAAACATGAAGAGT	1140
Db	1081	GCAAGAGGAGAGGGCAGAGCGATGAGTACTCGAGTAGTGAGAAACATGAAGAGT	1140
Qy	1141	GATAGTCAAGTCTCAAAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTC	1200
Db	1141	GATAGTCAAGTCTCAAGAGCAGCTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTC	1200
Qy	1201	AAAGAAAGGCTCCGAAGAGAGGGAGATATACCAAACCAATCAACTTGAGAGAGCGGA	1260
Db	1201	AAAGAAAGGCTCCGAAGAGAGGGAGATATACCAAACCAATCAACTTGAGAGAGCGGA	1260
Qy	1261	GCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAGAACCC	1320
Db	1261	GCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAGAACCC	1320
Qy	1321	CCAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTGAT	1380
Db	1321	CCAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTGAT	1380
Qy	1381	GCTCCACACTTCAACTCAAAAGGCCATGTTATCGTCTCGTCAACAAAGGAACTGGA	1440
Db	1381	GCTCCACACTTCAACTCAAAAGGCCATGTTATCGTCTCGTCAACAAAGGAACTGGA	1440
Qy	1441	CTTTGAACCTCGTGGCTGTAAAGAAAGACAAACAGAGAGGGAGCGGGAGAGAGGA	1500
Db	1441	CTTTGAACCTCGTGGCTGTAAAGAAAGACAAACAGAGAGGGAGCGGGAGAGAGGA	1500
Qy	1501	GGAGGAGAGAGAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCAGGTT	1560
Db	1501	GGAGGAGAGAGAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCAGGTT	1560
Qy	1561	GAAAGGAGGCGATGTGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC	1620
Db	1561	GAAAGGAGGCGATGTGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC	1620
Qy	1621	CGAATCTCATCTCGTTCGGCTTCGGTATCAACCGTGAACCAACAGATCTTCTCTGC	1680
Db	1621	CGAATCTCATCTCGTTCGGCTTCGGTATCAACCGTGAACCAACAGATCTTCTCTGC	1680
Qy	1681	AGGTGATTAAGGACAAATGTGTATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTTCC	1740
Db	1681	AGGTGATTAAGGACAAATGTGTATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTTCC	1740
Qy	1741	TGGGTCCGGTGAAACAGTTGAGAGGCTCATCAAAAACAGAGAGGAATCTCACCTTTGTGAG	1800
Db	1741	TGGGTCCGGTGAAACAGTTGAGAGGCTCATCAAAAACAGAGAGGAATCTCACCTTTGTGAG	1800
Qy	1801	TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGA	1860
Db	1801	TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGA	1860

Qy 1561 GAAGGAAAGCGATGTGTTTCATCATGCGCAGAGCTCATCCAGTAGCCATCAAGCTTCCTC 1620
Db 1561 GAAGGAAAGCGATGTGTTTCATCATGCGCAGAGCTCATCCAGTAGCCATCAAGCTTCCTC 1620
Qy 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACACAGAAATCTTCTTGC 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACACAGAAATCTTCTTGC 1680
Qy 1681 AGGTGATAAGACAAATGTGATAGACCAAGATAGAGAACCAAGCAAGCAATTTAGCATTCCTC 1740
Db 1681 AGGTGATAAGACAAATGTGATAGACCAAGATAGAGAACCAAGCAAGCAATTTAGCATTCCTC 1740
Qy 1741 TGGGTCCGGGTGAACAAAGTTCGAGAGCTCATCAAAACCAAGCAAGCAATTTAGCATTCCTC 1800
Db 1741 TGGGTCCGGGTGAACAAAGTTCGAGAGCTCATCAAAACCAAGCAAGCAATTTAGCATTCCTC 1800
Qy 1801 TGCTCGTCCCTCAATCTCAATCTCCGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Db 1801 TGCTCGTCCCTCAATCTCAATCTCCGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Qy 1861 AGAGGATCAAGAGAGAGAAACCAAGGAGGAGGGTCCACTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGAGAGAAACCAAGGAGGAGGGTCCACTCTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAAGATCACGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAAGATCACGCTTTTGTACT 1980
Qy 1981 CTACTATCCAAAACTTATCAATAATAAAAAAGTTTGTGGTTTGTCTCC 2032
Db 1981 CTACTATCCAAAACTTATCAATAATAAAAAAGTTTGTGGTTTGTCTCC 2032

RESULT 5

US-10-728-051-1
; Sequence 1, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10/728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-728-051-1

Query Match 100.0%; Score 2032; DB 20; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATAATCATATATATTCATCAATCATCTATATATAGTAGCAGAGCAATGAGGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGAGCAATGAGGGAG 60
Qy 61 GGTTCCTCCATGATGCTGTGTAGGATCCTTGTCTCTGCTTCAGTTCTGCAACGCA 120
Db 61 GGTTCCTCCATGATGCTGTGTAGGATCCTTGTCTCTGCTTCAGTTCTGCAACGCA 120
Qy 121 TGCCAAGTTCATCCCTTACCAGAGAAAAACAGAGAACCCCTTGCGCCAGAGGTGCTCCA 180

Db 121 TGCCAAGTTCATCCCTTACCAGAGAAAAACAGAGAACCCCTTGCGCCAGAGGTGCTCCA 180
Qy 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGTGCACCAA 240
Db 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGTGCACCAA 240
Qy 241 GCTCGAGTATGATCTCTCGTTGTGTATGATCTCTCGAGACACATCTGGACCAACCAACCA 300
Db 241 GCTCGAGTATGATCTCTCGTTGTGTATGATCTCTCGAGNACACATCTGGACCAACCAACCA 300
Qy 301 ACCTTCCCTTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATGACCG 360
Db 301 ACCTTCCCTTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATGACCG 360
Qy 361 CGGTCAACCCCGNAGAGAGAGGAGCGGATGGGACCAAGCTGGACCGAGGGAGCGTGA 420
Db 361 CGGTCAACCCCGNAGAGAGAGGAGCGGATGGGACCAAGCTGGACCGAGGGAGCGTGA 420
Qy 421 AAGAGAAAGAGACTGGAGACCAACCAAGAAAGATTGGAGGCGACCAAGTCAATCAGCAGCC 480
Db 421 AAGAGAAAGAGACTGGAGACCAACCAAGAAAGATTGGAGGCGACCAAGTCAATCAGCAGCC 480
Qy 481 ACAGAAAAATAAGGCCCCGAAAGAGAAAGAGAAAGAGAGTGGGAAACACACAGGTAGCCA 540
Db 481 ACAGAAAAATAAGGCCCCGAAAGAGAGAAAGAGAGTGGGAAACACACAGGTAGCCA 540
Qy 541 TGTGAGGAGAAACATCTCGNACCAACCTTCTACTTCCGTCGAGGGGTTTAGCAC 600
Db 541 TGTGAGGAGAAACATCTCGNACCAACCTTCTACTTCCGTCGAGGGGTTTAGCAC 600
Qy 601 CCCTACGGGAACCAAAACGCTAGGATCGGGTCTCTGACAGAGTTTGAACAAAGTCAAG 660
Db 601 CCCTACGGGAACCAAAACGCTAGGATCGGGTCTCTGACAGAGTTTGAACAAAGTCAAG 660
Qy 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTCAGATCGAGGCCAAACCTTAACACTCT 720
Db 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTCAGATCGAGGCCAAACCTTAACACTCT 720
Qy 721 TGTTCCTCCNAGCAGCTGATGCTGATAACATCTTGTATCCACAGGCGCAAGCCAC 780
Db 721 TGTTCCTCCNAGCAGCTGATGCTGATAACATCTTGTATCCACAGGCGCAAGCCAC 780
Qy 781 CGTGACGCTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCCATGCAC 840
Db 781 CGTGACGCTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCCATGCAC 840
Qy 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGGCATGACAAACAGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGGCATGACAAACAGAACTCAG 900
Qy 901 AGTAGCTTAAATCTCCATGCCGTTAACACACCCGGCCAGTTTGAAGATTTCTTCCGGC 960
Db 901 AGTAGCTTAAATCTCCATGCCGTTAACACACCCGGCCAGTTTGAAGATTTCTTCCGGC 960
Qy 961 GAGCAGCCGAGACCAATCATCTTCTTTCAGGGCTTCAGCAGGAATACGTTGAGGCGC 1020
Db 961 GAGCAGCCGAGACCAATCATCTTCTTTCAGGGCTTCAGCAGGAATACGTTGAGGCGC 1020
Qy 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA 1080
Db 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGGCGAGGCGATCGAGTACTCGAGGTAGTGAGAACAAATGAAGGT 1140
Db 1081 GCAAGAGGAGAGGGCGAGGCGATCGAGTACTCGAGGTAGTGAGAACAAATGAAGGT 1140
Qy 1141 GATAGTCAAAAGTCTCAAAGAGGACGTTGAAGAACTTACTTAAGCAAGTAAATCCGCTC 1200
Db 1141 GATAGTCAAAAGTCTCAAAGAGGACGTTGAAGAACTTACTTAAGCAAGTAAATCCGCTC 1200
Qy 1201 AAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCCCAATCAATTTGAGAGAGCGCA 1260

Db 961 GAGCAGCGGAGCAATCATCTTTCGAGGGCTTCAGCAGGAATACGTTGGAGGCCG 1020
Qy 1021 CTTCAATCCGGAATTCATGAGATACGAGGGTCTCTAGAGAGATCCAGAGGTGA 1080
Db 1021 CTTCAATCCGGAATTCATGAGATACGAGGGTCTCTAGAGAGATCCAGAGGTGA 1080
Qy 1081 GCAAGAGAGAGAGCGGAGAGCGCATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT 1140
Db 1081 GCAAGAGAGAGAGCGGAGAGCGCATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT 1140
Qy 1141 GATAGTCAAAGTGTCAAAGAGACGTTGAAGAACTTACTAAGCACGCTAAATCCGCTC 1200
Db 1141 GATAGTCAAAGTGTCAAAGAGACGTTGAAGAACTTACTAAGCACGCTAAATCCGCTC 1200
Qy 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACCACCCCAATCACTTGAGAGAGCGGA 1260
Db 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACCACCCCAATCACTTGAGAGAGCGGA 1260
Qy 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGTGAAGCCAGACAAGAAACCC 1320
Db 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGTGAAGCCAGACAAGAAACCC 1320
Qy 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAGAGCTTTGAT 1380
Qy 1381 GCTCCACACATTCACATCAAAGGCCATGGTTATCGTGTGTCAACAAAGGAACCTGAAA 1440
Db 1381 GCTCCACACATTCACATCAAAGGCCATGGTTATCGTGTGTCAACAAAGGAACCTGAAA 1440
Qy 1441 CTTTGAATCTGCTGTGAAGAAAGAGCAACACAGAGGGGAGCGGGAAGAGAGGA 1500
Db 1441 CTTTGAATCTGCTGTGAAGAAAGAGCAACACAGAGGGGAGCGGGAAGAGAGGA 1500
Qy 1501 GGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGTT 1560
Db 1501 GGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGTT 1560
Qy 1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
Db 1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
Qy 1621 CGAACTCCATCTGCTTGCCTTCGATCAACGCTGAAAAACACACAGAAATCTTCTTTC 1680
Db 1621 CGAACTCCATCTGCTTGCCTTCGATCAACGCTGAAAAACACACAGAAATCTTCTTTC 1680
Qy 1681 AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAGCAAGCGAAGGATTTAGCATCCC 1740
Db 1681 AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAGCAAGCGAAGGATTTAGCATCCC 1740
Qy 1741 TGGGTGGGTGAAACAAGTTGAGAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTGGGTGAAACAAGTTGAGAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGCTCAATCTCAATCTCAATCTCGCTGCTCCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGCTCGCTCAATCTCAATCTCAATCTCGCTGCTCCTGAGAAAGAGTCTCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAAAAACCAAGGAGGGAAGGGTCCACTCTCTTCAATTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAAAAACCAAGGAGGGAAGGGTCCACTCTCTTCAATTTGAAGGC 1920
Qy 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATAATAAGATCAGCGTTTGTACT 1980
Db 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATAATAAGATCAGCGTTTGTACT 1980
Qy 1981 CTACTATCCAAAACTTATCAATAAATAAAGCGTTTGTGCTTCTCC 2032
Db 1981 CTACTATCCAAAACTTATCAATAAATAAAGCGTTTGTGCTTCTCC 2032

US-10-958-324-5
; Sequence 5, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VQUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-10-958-324-5

Query Match 100.0%; Score 2032; DB 21; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAG 60
Qy 61 GGTTCCTCCACTGATGCTGTGTAGGGATCCTTGTCTGGCTTCAGTTCTGCAACGCA 120
Db 61 GGTTCCTCCACTGATGCTGTGTAGGGATCCTTGTCTGGCTTCAGTTCTGCAACGCA 120
Qy 121 TGCAGAGTATCATCACTTACCAGAAAGAAAAACAGAGAAACCCCTGCGCCACAGAGTGCCTCCA 180
Db 121 TGCAGAGTATCATCACTTACCAGAAAGAAAAACAGAGAAACCCCTGCGCCACAGAGTGCCTCCA 180
Qy 181 GAGTTGTCAACAGAAACCGGATGACTTGAAGCAAAAAGGATCGAGTCTCGCTGCACAA 240
Db 181 GAGTTGTCAACAGAAACCGGATGACTTGAAGCAAAAAGGATCGAGTCTCGCTGCACAA 240
Qy 241 GCTCGAGTATGATCCTCGTTGTCTATGATCTCGAGGACACACTGGCACCAACCAACCA 300
Db 241 GCTCGAGTATGATCCTCGTTGTCTATGATCTCGAGGACACACTGGCACCAACCAACCA 300
Qy 301 ACCTTCCCTCCAGGGGAGCGGACACGTGGCCGCCAAACCCGAGAGACTTACGATGATGACCG 360
Db 301 ACCTTCCCTCCAGGGGAGCGGACACGTGGCCGCCAAACCCGAGAGACTTACGATGATGACCG 360
Qy 361 CCCTCAACCCCGAAGAGAGAGGAGCGGATGGGACCAAGTTCGAGCGGAGCGGTGA 420
Db 361 CCCTCAACCCCGAAGAGAGAGGAGCGGATGGGACCAAGTTCGAGCGGAGCGGTGA 420
Qy 421 AAGAGAGAGACTTGGAGACAAACAGAGAGAGTTCGAGCGGACCAAGTTCATCAGCAGCC 480
Db 421 AAGAGAGAGACTTGGAGACAAACAGAGAGAGTTCGAGCGGACCAAGTTCATCAGCAGCC 480
Qy 481 ACCGAAAAATAAGGCCCGAAG 540
Db 481 ACCGAAAAATAAGGCCCGAAG 540
Qy 541 TGTGAGGAGAGAAACATCTTCGGAACAAACCCCTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Db 541 TGTGAGGAGAGAAACATCTTCGGAACAAACCCCTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Qy 601 CCCTACCGGAAACCAAAACCGTAGGATCCGGTTCCTGACAGGTTTGACCAAGGTCAG 660
Db 601 CCCTACCGGAAACCAAAACCGTAGGATCCGGTTCCTGACAGGTTTGACCAAGGTCAG 660

Db 663 CGAAGATG-----AACCATTTCAACTTTGAGAAGCGCAACCCCATCTATTTCACAA 713
Qy 1279 CTTTGGGAAGTTATTTGAGTGAAGCCAGACAAGAAACCCCGAGTTTCAGGACCTGGA 1338
Db 714 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGGACTTGG 770
Qy 1339 CATGATGCTACCTGCTGAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTC 1398
Db 771 TAATCTCTCAGTCTGTGATATCAAGAAGAGCTTCTTCTTACACACTTCAATTC 830
Qy 1399 AAAGCCATGTTATCTGCTGTCACAAAGAACTGGAACCTTGAATCTGCTGCTG 1458
Db 831 AAAGCCATGTTATCTGCTGTCACAAAGAACTGGAACCTTGAATCTGCTGCTG 890
Qy 1459 AAGAAAGACAAACAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
Db 891 TAAAGAAACAAACAG 933
Qy 1519 GGAGGGAAGTAACAGAGAGGTCGTAGGTACACAGCGAGGTTGAAGGAAGCGCATGTGTT 1578
Db 934 -----GAAGTGCAAGGTACAGAGCTGAATTTGCTGCAAGACGATGTAAT 977
Qy 1579 CATCATGCCAGACTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGG 1638
Db 978 TGTAAATCCAGAGCTTATCCATTTGCTGCTCAACGCTACCTCAAAACCTCAATTTCTCTGC 1037
Qy 1639 CTTTGGTATCAATCTGAGAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1698
Db 1038 TTTTGGTATCAATCTGAGAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
Qy 1699 GATGACACAGATAGAGAAGCAAGCAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGT 1758
Db 1098 GGTAAAGCAGATAGAGAAGCAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGT 1157
Qy 1759 TGAGAAGCTCATCAAAACAGAGAGATCTCACTTTGTGAGTGCTCGTCTCA 1812
Db 1158 TGAGAGGCTATTAAGAAGCAGAGGGAATCCTACTTTGTTGATGCTCAGGCTCA 1211

RESULT 11

US-09-758-652-3

; Sequence 3, Application US/09758652

; Patent No. US2001001137A1

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WINDOWS 95

; SOFTWARE: MICROSOFT WORD 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/758,652

; FILING DATE: 11-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/019,940

; FILING DATE: JUNE 14, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: LYNNE M. CHRISTENBURY

; REGISTRATION NUMBER: 30,971

; REFERENCE/DOCKET NUMBER: BB-1071-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-5481

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1320 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

; US-09-758-652-3

Query Match 20.0%; Score 406.8; DB 9; Length 1320;

Best Local Similarity 62.1%; Pred. No. 2.4e-115;

Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;

Qy 562 GAACAACCCCTTTCTACTTT---CCCGTCAAGCGCGTTTAGACCCCGTTCAGCGTTCAGGAAACCAAAA 618
Db 90 GAATAACCCCTTTCTACTTTTGAAGCTCTAACAGCTTCCAAACTCTCTTTTGAGAACCAAAA 149
Qy 619 CGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAGGTCAAGGCAGTTTCAGAAATCTCCA 678
Db 150 CGTTCGATTCGTCTCTCCAGAGATTCAACAAACGCTCCCACTTGGAAACCTTCG 209
Qy 679 GAATCACCGTATTGTGCAGATCGAGGCCAAACCTAAACACTCTTTGTTCTTCCCAAGCACGC 738
Db 210 AGACTACCGATTGTCCAGTTTCAGTCAAAACCCACACAATCTTCTTCCCCACCATGC 269
Qy 739 TGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCCACGTCGACCGTAGCAATGG 798
Db 270 TGACGCGGATTTCTCTCTTTGCTTAGCGGGAGAGCCATACTTACCTTTGGTGAACA 329
Qy 799 CAATAACAGAAAGCTTTTAATCTTCAGAGGCGCATGCACTCAGATCCCATCCGCTTT 858
Db 330 CGACGACAGAGACTCTTCAACCTTCACCTCGCGATGCGCAGAGAAATCCAGCTGAAC 389
Qy 859 CATTTCTCATCTTGAAACCGCATGACAAACAGAACCTCAGAGTAGCTAAAAATCTCCAT 918
Db 390 CACTTACTATTGGTTAAACCTCAGACACCAAGATCTCAAAATATCAAACTTGCCAT 449
Qy 919 GCCCGTTAACAACCCCGGCAGTTTGAAGGATTTCTTCCCGGCGAGCAGCGCAGACCAATC 978
Db 450 ACCCGTCAACAAACCTGGCAGATATGATGATTTCTTCTTATCTAGCACTCAAGCCCAACA 509
Qy 979 ATCTTACTTCAGGGCTTCAGAGGATAGCTTTGGAGGCGCGCTTCAATGGGAATTC 1038
Db 510 GTCTTACTTGAAGGCTTCAGCCATTAATATTCAGAGACCTCTCTTCCATAGCGAATTCGA 569
Qy 1039 TGAGATACGGAGGTCGTGTTAGAAGAGAAATGCAGGAGGTGAGCAAGAGAGAGAGGCA 1098
Db 570 GGAGATAAACAGGTTTGTGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621
Qy 1099 GAGGCGATGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTAGTAGTCAAGTGTCAAA 1158
Db 622 -----GGAGTGTGTTGAATCTCTCAA 644
Qy 1159 GGAGCAGCTTGAAGAACTTACTAAGCACGCTTAATCGTCTCAAGAAAGGCTCCGAAGA 1218
Db 645 GGAAACAAATTCGGCACTGAGCAGAGCTGCAAAATCTAGTTCAAGGAAACCAATTCCTC 704
Qy 1219 AGAGGGAGATATCAACCAACCCAATCAACTTTGAGAGAGGCGAGCCCGATCTTTTAAACA 1278
Db 705 CGAAGATG-----AACCATTTCACTTTGAGAGCGCAACCCCATCTATTCCAAACA 755
Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCCCGAGTTTCAGGACCTGGA 1338
Db 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGGACTTGA 812
Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAGAGAGGCTTTGATGCTCCACACTTCAACTC 1398
Db 813 TATCTCTCTCAGTCTGTGTGATATCAACGAAGGAGCTCTTCTTCTTACACACTTCAATTC 872

QY 1399 AAAGGCCATGTTATCGTCTCAACAAAGGAACCTGGAACCTTGAACCTCGTGGCTGT 1458
 DB 873 AAAGGCCATGTTATCGTCTCAACAAAGGAACCTGGAACCTTGAACCTCGTGGCTGT 932
 QY 1459 AAAGAAAAGACAAACAGAGGGGACGGCGGAAGAGAGGAGGACGAAGACGAAGAAGA 1518
 DB 933 TAAAGAAACAAACAGAGGACGAACAGGAGGAACTTTG----- 975
 QY 1519 GGAGGAAGTAACAGAGAGGTCGTAGGTACACAGCGAGGTTGAAGGAGGCGATGTGT 1578
 DB 976 -----GAAGTCAAGAGGTACAGAGCTGAAATGTCTGAAGACGATGTATT 1019
 QY 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACCTCTCCCGAATCCATCTCGCTGG 1638
 DB 1020 TGTAAATCCAGAGCTTATCCATTTGTCTCAACCTACCTCAACCTCAATTTCCCTTGC 1079
 QY 1639 CTTCCGTATCAACCGCTGAAAACCAACACAGATCTTCTTCAGGTGATAAGGACAATGT 1698
 DB 1080 TTTTGGTATCAATGCTGAGAACCAACAGAGGAACCTTCTTCAGCGGAGAAACAATGT 1139
 QY 1699 GATAGACAGATAGAGAAAGCAAGCAAGGATTTAGCATTTCCCTGGGTGGGTGAACAAGT 1758
 DB 1140 GGTAAAGGAGATAGAAAGCAAGTGCAGAGCTTCCGTTCCTGGGTCTGCACAAGATGT 1199
 QY 1759 TGAGAGCTCATCAAAAACCAAGGAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
 DB 1200 TGAGAGCTTATTAAGAGAGGAGGAATCTTCTTTGTGATGCTCAGCCTCA 1253

RESULT 12

US-10-245-227B-3
 ; Sequence 3, Application US/10245227B
 ; Publication No. US20030200558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rapp, William
 ; APPLICANT: Peng, Jiekin
 ; APPLICANT: Nadig, Gautham
 ; APPLICANT: Venkatesh, T
 ; TITLE OF INVENTION: ENHANCED PROTEINS AND METHODS FOR THEIR USE
 ; FILE REFERENCE: REN-00-087-US
 ; CURRENT APPLICATION NUMBER: US/10/245,227B
 ; CURRENT FILING DATE: 2002-09-17
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1320
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-245-227B-3

Query Match 20.0%; Score 406.8; DB 17; Length 1320;
 Best Local Similarity 62.1%; Pred. No. 2.4e-115;
 Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
 QY 562 GAACAACTTCTACTT---CCCGTCAAGCGGTTTAGCACCCGCTACGGGACCAAAA 618
 DB 90 GAATAACCTTCTACTTCTGAGAAGCTCTAACAGCTTCCAACTCTCTTTGAGAACCAAAA 149
 QY 619 CGGTAGGATCCGGGTCCTGACAGAGTTTGACCAAGGTTCAAGGCAAGTTTCAGAACTCCA 678
 DB 150 CGGTGCAATTCGTCTCTCCGAGAGATTCAACAGCTCCCCACACACTTGAGAACCTTCG 209
 QY 679 GAATCACCAGTATGTGACAGATCGAGGCCAAACCTTAACACTCTTTGTTCTCCCAAGCAGC 738
 DB 210 AGACTACCGGATTTGCCAGTTTTCAGTCAAAACCCCAACAATCTCTTCTCCCCCACTGC 269
 QY 739 TGATGCTCATTAACATCTTGTATTCACAGAGGCAAGCCAGCCGTGACCGTAGCAATGG 798
 DB 270 TGACGCCGATTTCTCTCTTGTCTTTAGCGGAGAGCCATCTTACCTTTGGTGAACAA 329
 QY 799 CAATAACAGAAAGAGCTTTAATCTTGACAGGGCCATGCTCAAGATCCCATCCGGTTT 858
 DB 330 CGACGACAGAGACTCTTAAACCTTTCACCTTGGCGATGCCAGAGATCCCGAGCTGGAAC 389

QY 859 CATTTCTTCACTTTTGAACCGCCATGACAAACGAACTCTAGAGTAGCTAAAAATCTCCAT 918
 DB 390 CACTTACTATTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 449
 QY 919 GCGGTTTAAACACACCCGCGCAGTTTGAAGATTTCTTCCCGCGAGAGCCGAGACCAATC 978
 DB 450 ACCGCTCAACAAACCTGCGAGATATGATGATTTCTTTATCTAGCATCTCAAGGCCAACA 509
 QY 979 ATCTTACTTTCAGGGCTTTCAGCAGGAATACGTTTGGAGCCGCTTCAATTCGGGAATTCAA 1038
 DB 510 GTCCTACTTTCAGGGCTTTCAGCAGGAATATTTCTAGAGACCTCTTCCATAGGAAATTCGA 569
 QY 1039 TGAGATACGAGGGTGTCTTTGAAGAGATGTCAGGAGGTGAGCAAGAGGAGAGAGGCA 1098
 DB 570 GGAGATAAACAGGGTTTGTGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
 QY 1099 GAGGCGATGGAGTACTCGGAGTAGTGAAACAATGAAGGAGGTGATGTCAAAGTGTCAAA 1158
 DB 622 -----GGAGTGATCTGGAACTCTCAAA 644
 QY 1159 GGAGCAGCTTGAAGAACTTACTAAGCAGCTTAAATCCGTCTCAAGAGAAAGCTCCGAAGA 1218
 DB 645 GGAACAAATTCGGCAACTGAGCAGACGCTGCCAAATCTAGTTTCAAGGAAACCACTTTCCTC 704
 QY 1219 AGAGGAGATATACCAACCAATCACTTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1278
 DB 705 CGAAGATG-----AACCAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 755
 QY 1279 CTTTGGGAGTTTATTTGAGGTGAAGCCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1338
 DB 756 CTTTGGGAAAGTTCTTTGAGATCACCCCTG---AGAAACCCCAAGCTTCCGGAGTTGGA 812
 QY 1339 CATGATGCTCACCTGTGTAGAGATCAAGAGAGGAGCTTTGATGTCTCCACACTTCAACTC 1398
 DB 813 TATCTCTCTCAGTTCTGTGGATATCAAGAGAGGAGCTTCTTCTTACCACTTCAATTC 872
 QY 1399 AAAGCCATGTTATCGTCTGTCGTCACAAAGAACTGGAACCTTGAACCTTCTGGGTGT 1458
 DB 873 AAAGCCATGTCATCTAGTATGATTAATGAAGGAGATGCAAACTTGAACCTTGTGGCAT 932
 QY 1459 AAGAAAGAGCAACAAACAGAGGGGCGCGGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518
 DB 933 TAAAGAAACAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
 QY 1519 GGAGGAAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAGGAGGAGGAGGAG 1578
 DB 976 -----GAAGTGCAAGGTACAGAGCTGMAATTGCTCAAGAGGAGTATT 1019
 QY 1579 CATATGCCAGAGCTCATTCAGTAGCCATCAACGCTTCTTCCGAACTCCCATCTGCTTGG 1638
 DB 1020 TGTAAATTCAGAGCTTATCCATTTGTCTCAACGCTACCTCAAAACCTCAATTTCTCTTGC 1079
 QY 1639 CTTCCGTTATCAACGCTCAAAACCAACCAAGATCTTCTTCCAGGTGATAGGACAAATGT 1698
 DB 1080 TTTTGGTATCAAGTCTGAGAAACCAAGAGAACTTCTTCCAGGCGGAGGAGGAGGAGGAGGAG 1139
 QY 1699 GATAGACAGATAGAGAAAGCAAGCAAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAGT 1758
 DB 1140 GGTAAAGCAGATAGAAAGACAGTGCAGAGAGCTTGGCTTCCCTGGGTCTGCACAAGATGT 1199
 QY 1759 TGAGAGCTCATCAAAACCAAGAGGAAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
 DB 1200 TGAGAGCTTATTAAGAGAGGAGGAAATCTTACTTTGTTGATGCTCAGCCTCA 1253

RESULT 13

US-10-684-651-3
 ; Sequence 3, Application US/10684651
 ; Publication No. US20040064858A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; GARY MICHAEL FADER

TITLE OF INVENTION: SUPPRESSION OF SPECIFIC CLASSES OF SOYBEAN SEED PROTEIN GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/684,651
FILING DATE: 14-Oct-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/758,652
FILING DATE: 11-Jan-2001
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14, 1996
ATTORNEY/AGENT INFORMATION:
NAME: LYNN M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 20.0%; Score 406.8; DB 18; Length 1320;
Best Local Similarity 62.1%; Pred. No. 2.4e-115;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
Qy 562 GAACAAACCTTTCTACTT---CCCGTCAGCGGTTTAGACCCGCTACGGGAACCAAAA 618
Db 90 GAATAACCTTTCTACTTTAGAGCTCTAACAGCTTCCAAACTCTCTTTGAGAACCAAAA 149
Qy 619 CGGTAGGATCCGGTCCCTGACAGGTTTGACCAAGGTTCAAGGAGTTTCAGATCTCCA 678
Db 150 CGTTGCGAATCGTCTCTCCAGAGATTCAACAAACGCTCCCAACAATTTGAGAACCTTCG 209
Qy 679 GAATCACCAGTATTGTGAGATCGAGGCAACCTTAACACTCTGTTCTTCCCAAGCAGC 738
Db 210 AGACTACCGGATTGTCCAGTTTCACTCAAAACCCCAACAACATCTCTTCCCCCACCATGC 269
Qy 739 TGATGCTGATAAATCCTTTGTTATCCAGCAAGGGAAGCCACCGTAGCCGTAGCAAAATGG 798
Db 270 TGACGCGGATTCTCTCTCTTTGTTCTTAGCGGAGAGCCATCTTACTTTGTTGTAACAA 329
Qy 799 CAATAACGAAGAGCTTTAATCTTGACGGGCCATCTCAGATCCCATCCGTTT 858
Db 330 CGACGACAGACTCTCTTACAACTTCCCTTCCCTGCGGAGAGCCATCTTCCAGTGGAAC 389
Qy 859 CATTTCTTACATCTTGAAACCCCACTGACCAACAGAACTCAGAGTAGTAAATCTCCAT 918
Db 390 CACTTACTATTGGTTAACTTACAGCACCAGCACTCAAAATCAATCAATTCGAT 449
Qy 919 GCCCGTTAAACACCCCGGCAATTTGAGGATTTCTTCCGGCGAGAGCCAGACCAATC 978
Db 450 ACCCGTCAACAAACCTGGCAGATATGATGATTTCTTTATCTATGCACTCAAGCCCAACA 509

Qy 979 ATCTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGGCGCCTTCAATCGGAATTCAA 1038
Db 510 GTCTACTTGCAGGGCTTCAGCCATTAATCTTAGAGACTCTCTTCCATAGCGAATTCGA 569
Qy 1039 TGAGATACGGAGGCTGCTGTTTGAAGAGAAATGAGAGGTTGAGCAAGAGAGAGAGGCA 1098
Db 570 GGAGATAAACAGGGTTTGTGTTGGAGAGGAAGAGGAGCAGAGCAGCAAG----- 621
Qy 1099 GAGGCGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGTAGTCAAGTGTCAAA 1158
Db 622 -----GGAGTAGTCTGGAATCTCAAA 644
Qy 1159 GGAGCAGCTTGAAGAACTTACTAAGCAGCGCTAAATCCGCTCTCAAGAAAGGCTCCGAAGA 1218
Db 645 GGACAAATTCGCAACTGAGCAGAGCTGCAAAATCTAGTTCAAGGAAACCAATTTCTC 704
Qy 1219 AGAGGAGATATCACCAACCAATCACTTGAAGAGAGGAGGAGCCGATCTTTCTAACAA 1278
Db 705 CGAAGATG-----AACCATTTCACTTGAAGCGCCCAACCCCATCTATTCCCACAA 755
Qy 1279 CTTTGGGAAGTTATTTGAGTGAAGCAGACAGAAAGACCCCGCAGCTTCAGGACCTGGA 1338
Db 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGGACTTGA 812
Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGAGGCTTTGATGCTCCCACTTCAACTC 1398
Db 813 TATCTCTCTCAGTTCTGTGATATCAACGAGAGAGCTCTTCTTCTACCACTTCAATTC 872
Qy 1399 AAAGGCCATGTTTATCGTGTCTCAACAAAGAACTTGAAGAACTTGAACCTCTGGCTGT 1458
Db 873 AAAGGCCATGATGATGATGATTAATGAAGAGAGATGCAAAACATTTGAACCTTGTGGCAT 932
Qy 1459 AAGAAAGAGCAACACAGAGGCGGCGGGAAGAGAGGAGGAGGAGCAAGACCAAGAGA 1518
Db 933 TAAAGAACCAACAGAGCAGAAACAGGAAGAGGAAACCTTTG----- 975
Qy 1519 CGAGGGAAGTAAACAGAGAGGTGCTAGGTAGGTACACAGCGAGGTTGAAGGAAGCGATGT 1578
Db 976 -----GAGTGCAAGGTACAGAGCTGAATTTGCTGTAAGACGATGTATT 1019
Qy 1579 CATCATGCGAGAGCTCATCCAGTAGCCATCAACGCTTCTCCGAATCTCATCTGCTGG 1638
Db 1020 TGTAAATTCAGCAGAGCTTATCCATTTGTGCTCAACGCTACCTCAAACTCAATTTCTTGC 1079
Qy 1639 CTTTGGTATCAACGCTGAAACCAACACAGAACTTCTTTCGAGGTGATAAGGACAATGT 1698
Db 1080 TTTTGGTATCAATGCTGAGAACCAACAGAGAACTTCTTTCGAGGCGGAGAAACAATGT 1139
Qy 1699 GATAGACCATAGAGAACCAAGAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAGT 1758
Db 1140 GGTAAAGCAGATAGAAAGCAAGTGCAGAGCTTGGTTCCTTGGGTCTGCACAAGATGT 1199
Qy 1759 TGAGAAGCTCATCAAAAACCAAGGAATCTCACTTTGTGAGTGTCTGCTCCTCA 1812
Db 1200 TGAGAGGCTATTAAAGAGCAGAGGGAATCCTACTTTGTTGATGCTCAGCCTCA 1253

RESULT 14
US-10-757-074-3
; Sequence 3, Application US/10757074
; Publication No. US20040139502A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON

STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/757,074
FILING DATE: 14-Jan-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/758,652
FILING DATE: 11-Jan-2001
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14,1996
ATTORNEY/AGENT INFORMATION:
NAME: LYNNE M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-757-074-3

Query Match 20.0%; Score 406.8; DB 19; Length 1320;
Best Local Similarity 62.1%; Pred. No. 2.4e-115;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
562 GAACACCCCTTCTACTT---CCGCTGAAGCGGTTTACACCGCTACGGGAACCAAAA 618
90 GAATAACCCCTTCTACTTTAGAGCTTAAAGCTTCAACAGCTTCCAAACTCTCTTTGAGAACCAAAA 149
619 CGGTAGATCGGGTCTGACAGGTTTGACCAAGGTCAGGAGTTCAGATCTCCA 678
150 CGTTCGCAATTCGTCTCTCCAGAGATCAACAAACGCTCCCAACCTTGAGAACCTTCG 209
679 GAATCACCGTATGTGCAGATCGAGGCCAAACCTTAAACACTCTGTCTTCTCCCAAGCAGC 738
210 AGACTACCGGATGTTCAGTTTCAGTCAAAACCCACACAACTCTTCTCCCCACCATGC 269
739 TGATGCTGATTAACATCTTGTATTATCCAGCAAGGCAAGCCACCGTAGCAAAATGG 798
270 TGACGCCGATTTCTCTCTTTGTCTTTAGCGGAGAGCCATCTTACCTTTGGTGAACAA 329
799 CAATAACAGAAGAGCTTATCTTTGACGAGGCGCATGCTCAGATCCCATCCGTTT 858
330 CGACGACAGAGACTCTTACAACTTCACTTCCGCGATGCCAGAGATCCAGCTGGAAC 389
859 CATTTCTTACATCTTGAACCCCATGACAAACAGAACTCAGAGTAGCTAAATCTCCAT 918
390 CACTTACTATTGGTTAACTCTACGACCAACAGATCTCAAAATATCAAACTTGCAT 449
919 GCCGTTTAACACACCCGCGAGTTTGGAGTATTTCTTCCGCGAGCAGCGAGCAATCA 978
450 ACCGCTCAACAACTGCGAGATATGATGTTTCTTATCTAGACACTCTTCCATAGCGAAATCGA 509
979 ATCTTACTTGAGGCTTTCAGCAGGAATACGTTGAGCGCCCTTCAATCGGGAATTCAA 1038
510 GTCTTACTTGCAAGGCTTTCAGCCATAATATCTAGAGACCTCTTCCATAGCGAAATCGA 569
1039 TGAGATACGAGGCTGCTGTTAGAGAGATGTCAGGAGGTGAGCAAGAGGAGAGGGCA 1098
570 GGAGATAACAGGGTTTGTGTTGAGAGGAAGAGAGAGAGGAGGAGCAAGAG----- 621

Qy 1099 GAGCGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGGTAGTATCAAGTGTCAA 1158
Db 622 -----GGAGTGTGATCGTGGAACTCTCAA 644
Qy 1159 GGAGCAGTTTGAAGAACTTAAAGCAGCGTAAATCGTCTCAAAGAAAGGCTCCGAAGA 1218
Db 645 GGAACAATTCGGCAACTGAGCAGACGTCGCAATCTAGTTCAAGGAAGAACCAATTCCTC 704
Qy 1219 AGAGGAGATATACCAACCAATCAACTTGAGAGAGGAGGAGCCGATCTTTTAACAA 1278
Db 705 CGAAGATG-----AACCAITCAACTTGAGAAGCGCAACCCCATCTATTCCAACAA 755
Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCAGACAGAAAGAACCCCGAGCTTCAGGACCTGGA 1338
Db 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGGACTTGA 812
Qy 1339 CATGATCTCACCTGTCTAGAGATCAAGAGGAGCTTTGATGCTCCACACTTCAACTC 1398
Db 813 TATCTTCTCAGTTCTGTGATATCAAGAGGAGCTCTTCTTCTACACACTTCAATTC 872
Qy 1399 AAAGCCCATGTTATCTGTCGTCAACAAAGGAACCTTGAACCTTCGAGCTGCTG 1458
Db 873 AAAGCCCATGTTATCTGATTAATGAAGGAGATGCAACACTTGAACCTTGTGGCAT 932
Qy 1459 AAGAAAGACCAACACAGAGGAGGCGCGGGAAGAGAGGAGGACGAGACGAGGAAGA 1518
Db 933 TAAAGAAACAAACAGAAAGCAGAAACAGGAAGAGGAACCTTTG----- 975
Qy 1519 GGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAAGCGGATGTT 1578
Db 976 -----GAAGTCAAGGTACAGAGCTGAATTTGCTGAAGACGATGATT 1019
Qy 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAATCCCATCTGCTGG 1638
Db 1020 TGTAAATTCAGCAGCTTATCCATTTGTCGTCAGCGTACCTCAAACTCAATTTCCCTTGC 1079
Qy 1639 CTTGCGTATCAACGCTGAAACCAACACAGAAATCTTCTTGCAGGTGATAGGACAATGT 1698
Db 1080 TTTTGGTATCAATGCTGAGAACCAACAGAGGAACTTCTTTCGAGGCGAGAAAGACAATGT 1139
Qy 1699 GATAGACCATAGAGAAAGCAAGAGGATTTAGCATTCCTTGGGTGGGTGAACAAGT 1758
Db 1140 GGTAAAGCATAGAAAGCAAGTGCAGAGGCTTGCCTTCCCTGGGTCTGCACAAGATGT 1199
Qy 1759 TGAGAAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
Db 1200 TGAGAGCTATTAAAGAGCAGAGGGAATCCTACTTTGTTGATGCTCAGCCTCA 1253

RESULT 15
US-10-757-155-3
; Sequence 3, Application US/10757155
; Publication No. US20040139503A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A

Search completed: August 24, 2005, 09:47:11
Job time : 1238.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 115.082 Seconds
(without alignments)
13658.075 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3562

Sequence: 1 aataacatatattcatc.....cggttgctggtgtttctcc 2032

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124355_29200/app_query.fasta_1.4757
-DB=A Geneseq 16Dec04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN_1_1_224 @runat_23082005_124355_29200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 16Dec04:*

1:	Geneseqp1980s:*
2:	Geneseqp1990s:*
3:	Geneseqp2000s:*
4:	Geneseqp2001s:*
5:	Geneseqp2002s:*
6:	Geneseqp2003as:*
7:	Geneseqp2003bs:*
8:	Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	92.3	626	2	Aay15244 Peanut al
2	3286	92.3	626	2	Aay25657 Peanut al
3	3286	92.3	626	6	Abus2412 Peanut Ar
4	3286	92.3	626	7	Adc34906 Peanut Ar
5	3282	92.1	626	4	Aau04706
6	3282	92.1	626	6	Abus2570 Peanut Ar
7	3280	92.1	626	6	Abus2568 Peanut Ar
8	3280	92.1	626	6	Abus2571 Peanut Ar
9	3280	92.1	626	6	Abus2574 Peanut Ar
10	3280	92.1	626	6	Abus2573 Peanut Ar

11	3279	92.1	626	8	ADM12053	Arachis h
12	3278	92.0	626	6	ABU52569	Peanut Ar
13	3278	92.0	626	6	ABU52572	Peanut al
14	3274	91.9	626	2	Aaw22150	Peanut al
15	3269.5	91.8	625	7	ADG27464	Peanut Ar
16	3251	91.3	626	8	ADO38271	Peanut al
17	3052	85.7	634	3	AAB33599	Modified
18	3052	85.7	634	4	Aau04709	Modified
19	3052	85.7	634	4	Aau05034	Modified
20	3041	85.4	614	2	Aaw22149	Peanut al
21	3041	85.4	614	2	Aaw62834	Arachis h
22	3041	85.4	614	6	ABU52413	Peanut Ar
23	3041	85.4	614	7	ADG27465	Peanut Ar
24	2130	59.8	415	2	AAY40913	Ara h 1 a
25	1281	36.0	605	2	Aaw62838	Glycine m
26	1281	36.0	605	7	ADG27564	Soybean B
27	1276	35.8	605	2	AAY40999	Soybean b
28	1190	33.4	559	5	ABG71270	Glycine m
29	1180.5	33.1	543	5	ABG71271	Glycine m
30	1165	32.7	417	5	ABG71269	Glycine m
31	1157	32.5	417	6	ABP97239	Mature be
32	1157	32.5	425	6	ABP97238	FLAG-tag
33	1156	32.5	439	6	ABP97235	Glycine m
34	1117.5	31.4	390	6	ABP97236	Glycine m
35	865.5	24.3	666	2	Aaw62829	Macadamia
36	863.5	24.2	625	2	Aaw62830	Macadamia
37	849.5	23.8	666	2	Aaw62828	Macadamia
38	811	22.8	566	2	AAR20181	Sequence
39	791	22.2	590	2	Aaw62832	Gossypium
40	764	21.4	371	2	AAY40914	Phaseolin
41	737.5	20.7	525	2	Aaw62831	Theobroma
42	654	18.4	366	6	ABU52503	Soybean B
43	629.5	17.7	593	2	Aaw62835	Zea mays
44	626	17.6	573	7	ADL18493	Maize glo
45	614	17.2	637	2	Aaw62837	Hordeum v

ALIGNMENTS

RESULT 1

AAV15244

ID AAV15244 standard; protein; 626 AA.

AC AAV15244;

XX

XX 17-OCT-2003 (revised)

DT 09-NOV-1999 (first entry)

XX

DE Peanut allergen, Ara h 1, amino acid sequence.

XX

KW allergy; immune response; transgenic; allergen; epitope;

KW immunoglobulin E; Ig E; binding site; peanut.

XX

OS Arachis hypogaea.

XX

PN WO9338978-A1.

XX

PD 05-AUG-1999.

XX

PF 29-JAN-1999; 99WO-US002031.

XX

PR 31-JAN-1998; 98US-0073283P.

PR

PR 13-FEB-1998; 98US-0074590P.

PR

PR 13-FEB-1998; 98US-0074624P.

PR

PR 13-FEB-1998; 98US-0074633P.

XX

XX 27-AUG-1998; 98US-00141220.

PA

(UYAR-) UNIV ARKANSAS.

PA

(UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

PA

(SOSI/) SOSIN H.

XX

PI Sosin H, Bannan GA, Burks AW, Sampson HA;

DR WPI; 1999-479189/40.
XX N-PSDB; AAZ06382.
XX Modified allergen with reduced IgE binding, useful for treating e.g.
PT allergies.
XX Disclosure; Page 35-37; 46pp; English.
XX This is the amino acid sequence of the Ara h 1 protein from Arachis
CC hypogaea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding
CC epitopes, four of which are immunodominant (AAI5247, AAI5249, AAI5250
CC and AAI5263). By modifying the IgE binding sites the ability of the
CC allergen to provoke an immune response is downregulated. The epitopes of
CC the IgE binding sites can therefore be modified in genetically engineered
CC plants and animals to elicit less of an allergic response. (Updated on 17
CC -Oct-2003 to standardise OS field)
XX Sequence 626 AA;
SQ
Alignment Scores:
Pred. No.: 1.4e-304 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-1 (1-2032) x AAI5244 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGTAGGATCTGTCTGCTTCAGTT 109
DB 1 MetArgLysArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACCATGCCAAGTATCATCCTTACCAAGAGAAACACAGAACCCCTGCCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCCAGATGTCAACAGGAACCGGATGACTTCAAGCAAAAGCATGCCAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGACCAAGTCGAGTATGATCTCTGTGTCTATGATCTTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTCGGCGGCCCAACCCCGGAGCTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGGACCACTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAGAGAGACTGGAGACAAACAGAGAGATGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGCAGCCAGGAAATAAGGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTyrGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACACCCCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCCGCTACGGGAACCAACCGTAGATCCGGGTCTCGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCGAGTTTCCAGATCTCCAGATCAACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAATGCAATAACAGAAAGAGCTTTAATCTTCCAGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCCATCCGTTTCTATCTTCTATCTTGAACGGCCATCACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTCAGAGT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGAGCGGAGAGCAATCATCTTCTTCTAGGGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATCCGGAATTCAATGAGATACCGAGGGTCTCTTGAAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGTGAGCAAGAGAGAGAGCGGAGAGGATCGAGTACTCGGAGTAGTCAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTAGTAGTCAAAAGTGTCAAAGGAGACGTTGAAGAACTTACTAAGCACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCAAGAAAGGCTCCGAGAGAGGAGGATATCACCAACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGCGAGCGCGATCTTTTAAACAACTTTGGAGTATTATTTAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCCTGCTGTGTAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTCAAAAGCGCTGTTTATCGTCTGCTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAATGGAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGG 1489
DB 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAG 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGTTTGAAGAGAGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGTCCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCCCTCCGAATCTCATCTCTGGCTTCGGTATCAACGCTGAAAAACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGGAGGTGATAGGACAAATGTGTATAGACCCAGATAGAGAGACGAGAGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATCTTCCCTGGGTGGTGAACAGTTTCAGAAAGCTCATCAAAACACAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTCTCTCTCTCTCTCT 1849

Db	581	HiePheValSerAlaIaIaProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
Qy	1850	TCTCCTCAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCACCTCCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGlnGlnGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGGCTTTTAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 2			
AAIY25657			
ID	AAIY25657	standard; protein; 626 AA.	
XX	XX	AAIY25657;	
XX	XX	17-OCT-2003 (revised)	
DT	30-SEP-1999	(first entry)	
XX	XX	Peanut allergen 1168391 Ara h I protein fragment.	
XX	XX	Major histocompatibility complex; class II; desensitising; human;	
KW	KW	allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;	
KW	KW	chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;	
KW	KW	screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;	
KW	KW	cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;	
KW	KW	mice; gerbil; vaccine; treatment; prevention; hypersensitivity; peanut.	
OS	OS	Arachis hypogaea.	
XX	XX	WO9934826-A1.	
XX	XX	15-JUL-1999.	
XX	XX	11-JAN-1999; 99WO-GB0000080.	
XX	XX	09-JAN-1998; 98GB-00000445.	
PR	PR	21-SEP-1998; 98GB-00020474.	
XX	XX	(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.	
PA	PA	Larche M, Kay AB;	
PI	PI	WPI; 1999-458255/38.	
XX	XX	Desensitizing patients to polypeptide allergens.	
DR	DR	Example 6; Page 70-71; 117pp; English.	
PT	PT	This invention describes a novel method of desensitizing a patient to a	
XX	XX	polypeptide allergen and comprises administering to the patient a peptide	
CC	CC	derived from the allergen where restriction to a MHC Class II molecule	
CC	CC	possessed by the patient can be demonstrated for the peptide and the	
CC	CC	peptide is able to induce a late phase response in an individual who	
CC	CC	possesses the MHC Class II molecule. The methods can be used for	
CC	CC	desensitising patients to allergens present in e.g. grass, tree and weed	
CC	CC	(including ragweed) pollens, fungi and moulds, foods, stinging insects,	
CC	CC	the chironomidae (non-biting midges), spiders and mites, housefly, fruit	
CC	CC	fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,	
CC	CC	non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of	
CC	CC	Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,	
CC	CC	sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to	
CC	CC	produce immunological vaccines which may be used to prevent and/or treat	
CC	CC	conditions involving hypersensitivity to allergens. This sequence	
CC	CC	represents a peanut (Arachis hypogaea) allergen 1168391 Ara h I. (Updated	
CC	CC	on 17-OCT-2003 to standardise OS field)	
XX	XX	Sequence 626 AA;	
SQ	SQ		
Alignment Scores:			
Pred. No.:	1.4e-304	Length:	626
Score:	3286.00	Matches:	626

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.25%	Indels:	0
DB:	2	Gaps:	0
US-10-728-323-1 (1-2032) x AAY25657 (1-626)			
QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTCTCTAGGATCTCTGTCTCGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGATGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGGCCAG	169
DB	21	SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln	40
QY	170	AGGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGTGCACCAAGCTCGAGTATGATCTCGTTGTCTATCATCTCGAGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCCTCCAGGGGAGCGACAGCTGGCCGCCAACCCGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCGCGTCCAAACCCGAGAGAGGAGGCGGATGGGAGCCAGCTGGACCG	409
DB	101	AspAspArgArgGlnProArgGluGluGlyArgGlyProAlaGlyPro	120
QY	410	AGGGAGCGTGAAGAGAAGAAGACTCGAGACAAACCAAGAGAAGATTGGAGGCCACCAAGT	469
DB	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer	140
QY	470	CATCAGCAGCCAGGAAATAGGCCCGAAGGAAGAGAGAGAAACAGAGTGGGAAACA	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTTAGCACCCGCTACCGGGAACCAAAACGGTAGGATCCGGCTCTCGAGAGGTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAGTCAAGGCGTTTCAGAAATCTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTGTCTTCCCAAGCAGCTGATGCTGATAACATCTCTGTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCGAAGCCACCGTACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACTCAGAAATCCATCCGGTTTCAATTTCTCATCTTTGAACCCGCCATGCAAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACCTCAGAGTAGTAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTGGAGT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGGCGAGCGACCAATCATCTACTTCAGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProIleSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGGAGCGCCCTTCAATCGGAATTCATATGATACGAGGGTGCTGTTAGAGAGAAAT	1069

Db	321	LeuGluAlaAlaPheAenAlaGluPheAenGluIleArgValLeuLeuGluGluAen	340
Qy	1070	GCAGGAGTGCAGCAAGAGGAGAGAGGCGCAGAGGCGATCGAGTACTCGGAGTAGTGAGAAC	1129
Db	341	AlaGlyGlyGluGluGluGluArgGlyGlnArgArgTrpSerThrArgSerGluAen	360
Qy	1130	AATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACGCT	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
Qy	1190	AAATCCGTCTCAAGAAAGGCTCCGAAGAAGAGGGAGATATCACCAACCCAACTCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
Qy	1250	AGAAAGCGCAGCCCGGATCTTCTAAACAACCTTTGGGAAGTATTATTGAGGTGAAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
Qy	1310	AAGAAGAACCCCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu	440
Qy	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATTCGTCGTCGCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
Qy	1430	GGAACCTCGAAACCTTTGAACCTCGTGGCTGTGAAGAAAGACAGCAACAAGAGGGGACGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgGly	480
Qy	1490	GAAGAAGAGGAGCGAAGACGAAAGAGAGGAGGGAATCAACAGAGAGTGCCTAGGTATC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	500
Qy	1550	ACAGCGAGTTGAAGGAAGCGATGCTGTTCAATCATGCCAGCAGCTCATCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
Qy	1610	AACGCTTCTCCGAACCTCCATCTGCTGGCTTCGGTATCAACGCTGAAAAACACCCACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAenAsnHisArg	540
Qy	1670	ATCTTCTCTGCAGGTGATAAGGACAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleaspGlnIleGluLysGlnAlaLysAsp	560
Qy	1730	TTAGCATTCCTGGGTGGGTGAACAAGTTGAGAGAGTCAATCAAAAAACAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
Qy	1790	CACTTTCTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
Qy	1850	TCTCCTCAGAAAGAGGATCAAGAGGAGGAAAAACCAAGAGGGAAGGTCCTCACTCTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAenGlnGlyGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 3			
ABU52412			
ID	ABU52412 standard; protein; 626 AA.		
XX	ABU52412;		
XX	AC		
XX	AC		
DT	10-MAR-2003 (first entry)		
XX	Peanut Ara h1 protein from cDNA clone P41b.		
DE	Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;		
KW			

anaphylactic food allergen; antiallergenic; vaccine; healing.	
Arachis hypogaea.	
WO200274250-A2.	
26-SEP-2002.	
18-MAR-2002; 2002WO-US009108.	
16-MAR-2001; 2001US-0276822P.	
19-MAR-2002; 2002US-00276822.	
(PANA-) PANACEA PHARM.	
Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G; Compadre CW, Connaughton C, Helm RW, King NE, Kopper RA, Maleki SJ; Rabjohn PA, Shin DS, Stanley JS;	
WPI; 2003-018765/01.	
N-PSDB; ABX70603.	
New modified anaphylactic food allergen, useful for preventing or treating allergic reactions associated with e.g. anaphylactic allergens.	
Claim 27; Fig 13; 300pp; English.	
The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IGG binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a peanut allergen (e.g. Ara h1, h2 or h3)	
Sequence 626 AA;	
Alignment Scores:	
Pred. No.:	1,4e-304
Score:	3286.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	92.25%
DB:	6
US-10-728-323-1 (1-2032) x ABUS2412 (1-626)	
Qy	50
Db	1
Qy	110
Db	21
Qy	170
Db	41
Qy	230
Db	61

QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGACAGTGGCCGCCAACCAGGAGACTAC 349
Db ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCCAAGAGAGAGAGAGCGCATGGCGACCGAGCTGGACCG 409
Db AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGGCGCAAGT 469
Db ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCAGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAAGTAGCCATGTCAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGAGAGGTTTGAC 649
Db ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGCTCAAGGAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAA 709
Db GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTCCCAAGCAGCTGTAGTCTGATAACATCTGTTATCCAGCAA 769
Db ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCCATCCGGTTTCATCTTCTACATCTTGAAACCGCATGACAAC 889
Db GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCCGCCAGTTTAGGAT 949
Db GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG 1009
Db PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATCGGAAATTCATAGATACGGAGGTGCTGTTAGAGAAAT 1069
Db LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAG 1129
Db AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTAGTAGTCAAGTGTCAAGAGGACGCTTCAAGAACTTACTAAGCACGCT 1189
Db AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AATCCGCTTCAAGAAAGCTCCGAAGAGAGAGAGAGATATCAACCAACCAATCAACTTG 1249
Db LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCATCTTCTTAACAATTTGGGAAGTTATTTGAGGTGAGCCAGAC 1309
Db ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTCGACATCATCTCACCTGTGTAGAGATCAAGAA 1369
Db LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440

QY 1370 GGAGCTTTGATGCTCCCACTTCACTCAAAAGGCGCATGTTATCGTCTCGTCAACAA 1429
Db GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGGCTGTAAGAAAAGAGCAACAACAGAGGAGCGCG 1489
Db GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArg 480
QY 1490 GAAAGAGAGGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
Db GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTCAAGAGAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCCCTCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACCAACACAGA 1669
Db AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGCAATGTGTATAGACCATAGAGAGAGAGAGAGAGAGAT 1729
Db IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGTCAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCT 1789
Db LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGAGAAAGAG 1849
Db HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAC 1927
Db IleLeuLysAlaPheAsn 626
RESULT 4
ADC34906
ID ADC34906 standard; protein; 626 AA.
XX
AC ADC34906;
XX
DT 18-DEC-2003 (first entry)
XX
DE Peanut allergen Ara h 1.
XX
KW allergen; antigen; hyporesponsive; desensitisation; immunomodulator;
KW gene therapy; peanut.
XX
OS Arachis hypogaea.
XX
PN WO2003047618-A2.
XX
PD 12-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-GB005548.
XX
PR 05-DEC-2001; 2001US-0338385P.
XX
PA (CIRC-) CIRCASSIA LTD.
XX
PI Larche M, Ledger PW;
XX
DR WPI; 2003-523267/49.
XX
PT Desensitizing an individual to a selected polypeptide antigen comprises
PT administering a composition containing polypeptide antigens in an amount
PT that generates a state of hyporesponsiveness to the antigen to allow
PT desensitization.

XX Disclosure; Page 40; 57pp; English.

PS The invention relates to a novel method for desensitising an individual

XX to a selected polypeptide antigen. The method comprises administering a

CC composition that contains polypeptide antigens in an amount that

CC generates in the individual a state of hyporesponsiveness to the antigen

CC to allow desensitisation to one or more polypeptide antigens. The method

CC of the invention has immunomodulator activity, and may have a use in gene

CC therapy. The composition and method are useful in manufacturing a

CC medicament for desensitising an individual to a selected polypeptide

CC antigen or for generating in the individual a state of hyporesponsiveness

CC to the antigen to allow desensitisation to one or more polypeptide

CC antigens. The present sequence is used in the exemplification of the

XX invention.

SQ Sequence 626 AA;

Alignment Scores:

Pred. No.:	1..4e-304	Length:	626
Score:	3286.00	Matches:	626
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.25%	Indels:	0
DB:	7	Gaps:	0

US-10-728-323-1 (1-2032) x ADC34906 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCTGATGCTGTTGGTAGGGATCCTTGCTCGCTTCAGTT 109

DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20

QY 110 TCTGCAACCCATGCCAAGTCATCACCTTACCAAGAGAAACAGAGAACCCCTGCCCCAG 169

DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGCTTGAAGCAAAAGCATCGCAGTCT 229

DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGCACCAAGCTCGAGTATGATCTGTTGTTGTTATGATCTCTCGAGGACACACTGGC 289

DB 61 ArgCysThrLysLeuGluLysArgProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGTTTCCCTCCAGGGAGCGACACGTCGGCCGCCAACCCCGAGACTAC 349

DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGGACCACTGGACCG 409

DB 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120

QY 410 AGGGAGCTGAAGAGAGAACTGGAGACCAACCAAGAGAGATTGGAGGCGACCAAGT 469

DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140

QY 470 CATCAGACCCACGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529

DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnLysIleArgTyr 160

QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACCAACCCCTTCTACTTCCCGTCAAG 589

DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTAGCACCGCTACGGGAACCAACCGTAGATCCGGGTCTCTCAGAGGTTTGAC 649

DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGGTCAAGCGAGTTTCCAGAAATCCAGAAATCACCGTATTGTGCAGATCCGAGCCAAA 709

DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCGTGTGCTGTGATAACATCTTGTATTATCCAGCA 769

DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY 770 GGGCAAGCCACCGTACCGTATAGCAATGCGCAATAACAGAAAGAGCTTTAACTTTACGAG 829

DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTTCTACATCTTGAACCGCATCACAAAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTCAGAT 949

DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCCTCCCGCGAGCGAGCGAGACCAATCATCTTCTGAGGGCTTCAGCAGGATACG 1009

DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGAGAGCGCGCTTCAATGCCGAATTCATAGATACCGAGGGTCTCTGTAGAAAGAGAT 1069

DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340

QY 1070 GCAGAGGTGAGCAAGAGAGAGAGCGAGCGAGCGATCGAGTACTCTCGAGTAGTGAGAAC 1129

DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360

QY 1130 AATGAGAGAGTATGATCAGAGTGTCAAGAGAGAGCGAGTGTGAAGACTTACTTAAGCACCT 1189

DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AAATCCGCTCAAAGAAAGGCTCCGAGAGAGAGAGAGATATCACCAACCCCAATCACTTG 1249

DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

QY 1250 AGAGAGCGAGCGCGATCTTTCTAAACAACTTTGGAAAGTTATTGAGGTGAAGCCAGAC 1309

DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAGAGACCCCGCTTCAGAGCTCGACATGATGCTCCTCTGTAGAGATCAAGAA 1369

DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440

QY 1370 GGAGCTTTGATGCTCCCACTCAACTCAAGGCCATGGTTATCGTCTGCTCAACAAA 1429

DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460

QY 1430 GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGCAACACAGAGGGGACGGCG 1489

DB 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAGAGGAGGAG 1549

DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY 1550 ACAGCAGGTTGAAGAGAGCGATGCTGTTCATCATGCCAGCAGCTCATCCAGTACCCATC 1609

DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520

QY 1610 AACGCTTCTCCGAACTCCTATCTCTGCTTGGCTTCCGTTATCAACGCTGAAAAACACCA 1669

DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540

QY 1670 ATCTTCTCTGAGGTGATAGGCAATGTGTATAGACCAAGAGAGAGAGAGAGAGAGAG 1729

DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560

QY 1730 TTAGCATTTCCCTGGGTCGGGTGAAACAGTTGAGAGAGCTCATCAAAAACAGAGAGAACT 1789

DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTGCTCTCTCTGAGAAAGAG 1849

Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
 QY 1850 TCCTCTGAGAAAGAGATCAAGAGGAGAAACCAAGAGGAGGAGGTCCTCTTCA 1909
 Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927
 Db 621 IleLeuLysAlaPheAsn 626

RESULT 5
 AAU04706
 ID AAU04706 standard; protein; 626 AA.
 XX
 AC AAU04706;
 DT 23-OCT-2001 (first entry)
 XX
 DE Anaphylactic antigen Ara h 1.
 XX
 KW Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse.
 XX
 OS Mus sp.

EH Key Location/Qualifiers
 FT Misc-difference 285 /note= "Encoded by gta"
 FT
 XX WO200140264-A2.
 PN
 XX
 PD 07-JUN-2001.
 XX
 PF 06-DEC-2000; 2000WO-US033124.
 XX
 PR 06-DEC-1999; 99US-00455294.
 PR 23-JUN-2000; 2000US-0213765P.
 PR 27-SEP-2000; 2000US-0235797P.
 XX

XX (PANA-) PANACEA PHARM LLC.
 PA (UYAR-) UNIV ARKANSAS.
 PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
 PA
 PI Bannan CA, Burks WA, Caplan MJ, Sampson H, Sosin H;
 XX
 DR WPI; 2001-381378/40.
 DR N-PSDB; AAS08537.
 XX

XX Antigenic fragments useful for reducing anaphylactic risk and reducing
 PT the severity and/or number of allergic symptoms in individuals sensitive
 PT to antigens, have reduced ability to bind Immunoglobulin E.
 XX
 PS Claim 7; Fig 9; 100pp; English.

XX The sequence represents the amino acid sequence of anaphylactic antigen
 CC Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design
 CC antigenic peptides having a reduced ability to bind immunoglobulin E
 CC (IgE) as compared with the intact (A), or having a sequence substantially
 CC identical to a portion of sequence of an antigen that includes at least
 CC one IgE binding site, where at least one IgE binding site of the peptide
 CC is altered. The antigenic peptides are used in a composition which is
 CC useful for reducing risk or severity of allergic reaction to an antigen.
 CC This is done by identifying an individual at risk of allergic reaction to
 CC an antigen by identifying prior display of allergic symptoms when exposed
 CC to the antigen, or a familial relationship with an individual who
 CC previously displayed allergic symptoms when exposed to the antigen.
 CC Following this an antigen-specific IgE present on one or more mast cells
 CC or basophils in the individual's serum is identified. The individual is
 CC then contacted with a peptide corresponding to a portion of the antigen,
 CC which is selected, formulated, and delivered so that binding of the
 CC peptide to antigen-specific IgE is reduced as compared with IgE binding
 CC of intact antigen. The composition is also useful for treating and
 CC preventing allergic reactions

XX SQ Sequence 626 AA;
 Alignment Scores:
 Pred. No.: 3,38e-304 Length: 626
 Score: 3282.00 Matches: 625
 Percent Similarity: 99.84% Conservatives: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.14% Indels: 0
 DB: 4 Gaps: 0
 US-10-728-323-1 (1-2032) x AAU04706 (1-626)
 QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCCTTGCTCTGGCTTCAGTT 109
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGCATGCCCAAGTCATCCTTACCAGAGAAAACAGAGAACCCCTGCGCCAG 169
 Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTCCCTCCAGAGTTGTCAACAGAACCGGATGACTTCAAGCAAAAGCATGCGAGTCT 229
 Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
 Db 61 ArgCysThrLysLeuGluLysAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGAGGACACGCTGGCGCCCAACCCCGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgLysArgLysProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 409
 Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgGlyTrpGlyProAlaGlyPro 120
 QY 410 AGGAGCGGTGAAGAGAGAGAGACTTGGAGACAAACCAAGAGAAAGATTGGAGGCGACCAAGT 469
 Db 121 ArgGluArgGluArgGluLysAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGACGACCGCAAAATAGGCCCGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 529
 Db 141 HisGlnGlnProArgLysLysLeuArgGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGAGAGAAACATCTCGAAACAAACCTTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGTTTTCAGCCGCTACGGGAAACCAAAACGGTAGGATCCGGGTCTCTGAGAGGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgLysValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGATCAACCGTATTGTTCAGATCCAGGCCAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgLysValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTGTCTTCCCAAGCACCTGATGTGATAACATCCTTGTATTATCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTACCGTACGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCTTCTACATCTTCTACATCTTCTACATCTTCT 889
 Db 261 GlyHisAlaLeuArgGlyProSerGlyPheLysSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGGTTAAACACACCCGCGCAGTTTCAGAGT 949
 Db 281 GlnAsnLeuArgAlaAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

170 AGTGCTCCAGATTGTCACAGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db |||||
41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY CGTGTGACCAAGCTCGAGTATGATCTCTGTTGTCTATCATCTCTCGAGACACACTGGC 289
Db |||||
61 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTCTCCCTCCAGGGAGCGGACACAGTGGCGGCCCAACCCGAGACTAC 349
Db |||||
81 ThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCTCAACCCCGAAGAGAGAGAGCGCGATGGGGACCGAGCTGACCG 409
Db |||||
101 AspAspArgArgGlnProArgGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGCTCAAGAGAAAGACTCGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT 469
Db |||||
121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 140
QY 470 CATCAGACGCCGGAATAAGCCCGAAGAGAGAGAGAAAGAGCAAGTGGGGAACA 529
Db |||||
141 HisGlnProArgLysIleArgProGluGlyArgGluGlyGlnGlnTyrGlyThr 160
QY 530 CCAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAG 589
Db |||||
161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCCGCTACGGGAACCAACCGTAGGATCCGGTCTCGACAGGTTTGAC 649
Db |||||
181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db |||||
201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTTCCAGACGCTGATGCTGATAACATCTTGTATTTCAGCAA 769
Db |||||
221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCACAGCCCGTGCAGCTAGCAATGGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db |||||
241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCATCAGAAATCCCATCCGTTTCCATTTCCCTACATCTTGAAACCGCCATGACAAC 889
Db |||||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCCATCCCGCTTAAACACACCCCGGCGAGTTTGAAGAT 949
Db |||||
281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGCGAGCGCGAGACCAATCATCTCTCTTGCAGGCTTTCAGCAGGAATACG 1009
Db |||||
301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCCCGCTTCAATCGGAAATTCATAGATACGGAGGCTGCTTTAGAGAGAAT 1069
Db |||||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGACGAGGAGAGAGCGGCGAGCGGATGGAGTACTCGAGTAGTGAGAC 1129
Db |||||
341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTGTAGTCAAGGTGTCAAAGGACGACGTTTCAAGAACTTACTAGACGCT 1189
Db |||||
361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCAAGAAAGCTCCGAAAGAGAGGAGATATCACCACCAACCAATCAACTTG 1249
Db |||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

1250 AGAGAGGCGAGCCCGGATCTTTCTAACACTTTGGGAAGCTTATTTGAGGTGAGCCAGAC 1309
Db |||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGAGCTTCCAGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db |||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGTCCTCCACACTTCAACTCAAAGGCCATGGTTATCTCGTCGTCAACAA 1429
Db |||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGCGCGCG 1489
Db |||||
461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAGCGAAGACCAAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTAC 1549
Db |||||
481 GluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGGAAGCGGATGTTCATCATGCCAGAGCTCATCCAGTAGCCATC 1609
Db |||||
501 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAAACACACAGAG 1669
Db |||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATTAAGGACAATGTGATAGACAGATAGAGAAGCAAGCAAGGAT 1729
Db |||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAGAGCTCATCAAAAACAGAGGAATCT 1789
Db |||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTCAGTGTCTCTCAATCTCAATCTCAATCTCTCGTCTCTCGAGGAAGAG 1849
Db |||||
581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAGAGAGGATCAAGAGGAGGAAAAACCAAGAGGGAAGGCTCCACTCTTTCA 1909
Db |||||
601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db |||||
621 IleLeuLysAlaPheAsn 626
RESULT 7
ABU52568
ID ABU52568 standard; protein; 626 AA.
XX AC ABU52568;
XX DT 10-MAR-2003 (first entry)
XX DX Peanut Ara h1 mutant K32A.
XX DE Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
XX KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
XX KW wound healing.
XX OS Homo sapiens.
XX OS XX
XX PN WO200274250-A2.
XX XX
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX XX
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX XX

PA (PANA-) PANACEA PHARM.
 XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabjohn PA, Shin DS, Stanley JS;
 XX WPI; 2003-018765/01.
 DR New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 XX Example 5; Page; 300pp; English.
 PS The invention relates to a modified anaphylactic food allergen that has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulfide bond. The
 CC modification may also comprise mutation of the IGE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a food
 CC allergen, mutated to alter its IGE binding characteristics. Note: The
 CC present sequence is not shown in the specification but was created by the
 CC indexer using information provided in the specification
 XX Sequence 626 AA;
 SQ

Alignment Scores:
 Pred. No.: 5.25e-304 Length: 626
 Score: 3280.00 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.08% Indels: 0
 DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABUS2568 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGCTAGGATCCTGTCTGCTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAGAGAACCCCTGCGCCAG 169
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 40
 QY 170 AGGTGCTCCAGATGTGTCACAGAGAACCGGATGACTTGAAGCAAAAGCATCGCAGTT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCAGATGATGCTCGTGTGCTATGATCCTCGAGGACACACTGGC 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCTCCCTCCAGGGAGCGGACACGTGGCGGCCAACCCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGAGCGCGATGGGGACCACTGGACCG 409
 DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGGAGCGTGAAGAGAGCAAGCACTGGAGCAACCAAGAGAGATTTGGAGGCGCAAGT 469
 DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 140

QY 470 CATCAGAGCCACGGAATAAAGGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGAACA 529
 DB 141 HisGlnGlnProArgLysLeuArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTTCTACTTCCCGTCAGG 589
 DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTACACCCGCTACGGGACCAAAACCGTAGGATCCGGTCTCCTCAGAGGTTTCAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGGCAGCTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCAGGCCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CTTAACTCTTGTGTTCTCCCAAGCACGCTGATGATGATGATGATGATGATGATGATGAT 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTCACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTCACGAG 829
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTATTTCTTACATCTTGAACCGCATCACAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTTCTTCCCGGAGCAGCCGAGACCAATCATCTTCTGAGGCTTCAGAGGAGTAACG 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGAGGCGCGCTTCAATGCGGAATTTCAATGAGATACGAGGGTGCTGTTGAAGAGAGAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGAGGTGAGCAAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1129
 DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGACGTTGAAGAACTTACTAAGCACGCT 1189
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGTCTCAAAGAAAGGCTCCGAAAGAGGAGAGATATCACCAACCCCAATCAACTTG 1249
 DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAGGCGAGCCCGATCTTCTTAAACAACTTGGGAAGTTATTGAGGTGAGGCCAGAC 1309
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAGAACCCCGCTCAGCTCAGACATGATGCTCACTGTGTAGAGATCAAGAA 1369
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
 QY 1370 GGAGCTTTGATGCTCCCACTCAACTCAAAGGCCATGGTTATCGTCTGCTCAACAAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAAGTGAACCTTGAATCTGTGCTGTAGAAAGAGCAACACAGAGGGGCGGG 1489
 DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
 QY 1490 GAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
 DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluGluValArgTyr 500

QY 1550 ACAGCGAGTTTGAAGGAGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTCAGGTGATAGGACATGTGATAGACCAGATAGAGCAAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCTCGGTCCGGTCCGAGTGAAGTTCAGAGCTCATCAAAACCAAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCTCTCGAAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 8
ID ABUS2571
AC ABUS2571 standard; protein; 626 AA.
AC ABUS2571;
XX
DT 10-MAR-2003 (first entry)
DE Peanut Ara h1 mutant R91A.
DE
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW munein; anaphylactic food allergen; anti-allergenic; vaccine;
KW wound healing.
XX
OS Homo sapiens.
XX
XX
XX WO200274250-A2.
XX
PD 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IgE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding

CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;
Alignment Scores:
Pred. No.: 5,25e-304 Length: 626
Score: 3280.00 Matches: 625
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 92.08% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABUS2571 (1-626)
QY 50 ATGAGAGGAGGGTTCCTCCACTGATGCTGTGCTAGGATCCTTGCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCCAAGTCATCACCTTACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCTCCAGAGTTGTCAACAGAGAACCGGATGACTTGAAGCAAAAGCATCGCAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGACCAAGCTCGAGTATGATCCTGTTGTGTATGATCCTCGAGGACACTGCG 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGACACGTCGGCCGCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAG 469
Db 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGACCCACGAGAAATATAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgGlyIleArgProGluGlyArgGluGlyGluGlnIleGluGlnIleGluGln 160
QY 530 CCAGGTAGCCATGTGAGGAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGTTTAGACCCGCTACGGAGAACAAACCGGTAGGATCCCGGTCTCGAGAGGTTTAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCAGTTTCAGAGATCTCCAGATCACCGTATGTCAGATCCAGGCCAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGlnAlaLys 220
QY 710 CTTAAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATGAACATCTTGTGTTATCAGCA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY	770	GGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG	829	QY	1850	TCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCTCCTTTCA	1909
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260	Db	601	SerProGluLysGluAspGlnGluGluGlnGlyLysGlyProLeuLeuSer	620
QY	830	GGCCATGCACTCAGAATCCCATCCGGTTTCATTCCTACATCTTGAACCGCCATGACAAC	889	QY	1910	ATTTTGAAGGCTTTTAAC	1927
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280	Db	621	IleLeuLysAlaPheAsn	626
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCGTAAACACACCCGCCAGTTGAGGAT	949	RESULT 9			
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300	ID	ABU52574	standard; protein; 626 AA.	
QY	950	TTCTTCCCGCCAGCAGCCGACCAATCATCTACTTCAGGGCTTCAGCAGGAATACG	1009	AC	ABU52574;		
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320	DT	10-MAR-2003	(first entry)	
QY	1010	TTGGAGCCGCTTCAATCGGAATTCATAGATACGAGGGTGCTGTGAAGAGAAAT	1069	XX	Peanut Ara h1 mutant R499A.		
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340	XX	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;		
QY	1070	GCAGGAGGTGACCAAGAGGAGGAGCGGACGAGCGATGAGTACTCGGAGTAGTGAGAC	1129	KW	mutein; anaphylactic food allergen; antiallergenic; vaccine;		
Db	341	AlaGlyGlyGlnGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAsn	360	KW	wound healing.		
QY	1130	AATGAAGGAGTGATAGTCAAAAGTGTCAAAAGGAGCACGTTTGAAGAACTTACTAAGCACGCT	1189	OS	Homo sapiens.		
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380	XX	WO200274250-A2.		
QY	1190	AAATCCGTCTCAAAAGAAAGCTCCGAAGAGGGAGATATACCAACCCAATCAACTTG	1249	PD	26-SEP-2002.		
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400	XX	18-MAR-2002; 2002WO-US0009108.		
QY	1250	AGAGAGGCGACCCGATCTTCTAACAACTTTGGGAAGTTATTTGAGTGAGCCAGAC	1309	PF	16-MAR-2001; 2001US-0276822P.		
Db	401	ArgGluGlyGluProAspLysSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420	PR	18-MAR-2002; 2002US-00276822.		
QY	1310	AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369	XX	(PANA-) PANACEA PHARM.		
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440	PA	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;		
QY	1370	GGAGCTTTGATGCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCGTCCTCAACAA	1429	XX	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460	PI	Rabjohn PA, Shin DS, Stanley JS;		
QY	1430	GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGACCAACAGAGGGGAGCGGG	1489	DR	WPI; 2003-018765/01.		
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480	XX	New modified anaphylactic food allergen, useful for preventing or		
QY	1490	GAAGAAGAGGAGGACGAAGACGAAGAGGAGGAAAGTAAACAGAGAGTGCGTAGGTAC	1549	PS	treating allergic reactions associated with e.g. anaphylactic allergens.		
Db	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500	CC	Example 5; Page; 300pp; English.		
QY	1550	ACAGCAGGTTGAAGAAGCGATGTTTCATCATCGCCAGCAGCTCATCCATGAGCCATC	1609	CC	The invention relates to a modified anaphylactic food allergen has an		
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520	CC	amino acid sequence that is substantially identical to that of natural		
QY	1610	AACGCTTCCTCCGAACTCCATCTGCTTGGCTTCGATCAAGCTCAAAACCAACACACAGA	1669	CC	anaphylactic food allergen, except for a cysteine residue that has been		
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540	CC	modified so that it cannot participate in the disulphide bond. The		
QY	1670	ATCTTCTTGCAGGTGATGAAGCAATGTGATAGACCAGATAGAAGCAACCGAAGGAT	1729	CC	modification may also comprise mutation of the IgE binding sites to		
Db	541	IlePheLeuAlaGlyAspLysAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560	CC	reduce allergenicity. Also included are: (1) a method of making a		
QY	1730	TTAGCATTCCTGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAACACAGAAGAAATCT	1789	CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding		
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580	CC	or for causing a site specific mutation in the modified anaphylactic food		
QY	1790	CACTTTGTGAGTGCCTCCTCAATCTCAATCTCAATCTCCGTCGTCCTCCAGAAAGAG	1849	CC	allergen; (3) a transgenic plant or animal expressing the modified		
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600	CC	anaphylactic food allergen; (4) a method of treating an individual by		
				XX	reducing the clinical response to a natural anaphylactic food allergen;		
				SQ	and an isolated fragment of peanut allergen Ara h 1. The modified		
				Sequence 626 AA;			
				Alignment Scores:			
				Pred. No.:	5.25e-304	Length:	626
				Score:	3280.00	Matches:	625
				Percent Similarity:	99.84%	Conservative:	0

Best Local Similarity: 99.84%		Mismatches: 1	
Query Match: 92.08%		Indels: 0	
DB: 6		Gaps: 0	
US-10-728-323-1 (1-2032) x ABU52574 (1-626)			
QY	50	ATCAGAGGAGGGTTCCTCCACTGATCTGCTGCTAGGATCCTTGTCTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGATGCCAAGTCATCACTTACAGAGAAACACAGAGAACCCCTCGCCACG	169
DB	21	SerAlaThrHisAlaLysSerProTyrGlnLysIleThrGluAsnProCysAlaGln	40
QY	170	AGTGCTCTCAGATTGTCAACAGAAACCGATGACTTGAAGCAAAAGGATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCAGATGATCTCTGTTGTCTATGATCTCTCGAGGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACACAGTGGCCGCCAACCCGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATACCGCTCAACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	409
DB	101	AspAspAspArgArgGlnProArgGlyGluGlyArgGlyArgGlyProAlaGlyPro	120
QY	410	AGGAGCGTGAAG	469
DB	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer	140
QY	470	CATCAGCAGCCAGCAAAATAAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGlyGluGlnGluTrpGlyThr	160
QY	530	CAAGGTAGCATGTGAGGAGAGAAACATCTCGGAACAAACCCCTTCTACTCTCCGCTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCGCTACCGGAGACCAAAACGGTAGGATCCGGTCTCGAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAAGTCAAGCATTTTCAGAAATCCAGAAATCCAGTATTGTGCAGATCCAGGCCAAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGGCAAGCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTTAACTCTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACTCAGAAATCCATCCGGTTTCATTTCTTCTACATCTTGAACCGCCATGACAAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACCTCAGATAGCTAAATCTCCATGCCCGTTTAAACACACCCGCCAGCTTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTACTTGCAGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGGCGCCCTTCAATCGGAAATTCATAGATACGGAGGGTCTGTTAGAGAGAAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
QY	1070	GCAGGAGGTGACCAAG	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
QY	1130	AATGAAGAGATGATAGTCAAAAGTGTCAAAAGGAGACGCTTGAAGAACTTACTAAGCACGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCAATCAACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGAGCCCGATCTTTTCTAACAACTTTTCGGAACTTATTTGAGGTCAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AGAAGAAGACCCCGAGCTTCAGACCTCGACATGATGCTCACCTGTGTGAGAGATCAAGA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCCACTTCAACTCAAAAGGCCATGTTTATCTGCTCGTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAATCGAAACCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAGCGAAGACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgAlaTyr	500
QY	1550	ACAGCGAGTTCGAGGAGGCGATGTTTCAATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGTTCTTCCGAATCTCATCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGAG	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTGCAGGTGATAAGGACAAATGTGTATAGACCAAGTACAGATAGAGAACGAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CATTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCTCGTGAAGAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlyLysGlu	600
QY	1850	TCTCTGTGAAGAGAGATCAAGAGAGGAGAAACCAAGAGGAGGAGGCTTCACCTCTTTCA	1909
DB	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC 1927	
DB	621	IleLeuLysAlaPheAsn 626	
RESULT 10			
ABU52573			
ID	ABU52573	standard; protein; 626 AA.	
XX	XX	ABU52573;	
XX	XX	10-MAR-2003 (first entry)	
XX	XX	Peanut Ara h1 mutant R109A.	
XX	XX	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
KW	KW	muten; anaphylactic food allergen; anti-allergenic; vaccine;	

KW wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO200274250-A2.
 XX
 PD 26-SEP-2002.
 XX
 XX 18-MAR-2002; 2002WO-US0009108.
 PF
 XX 16-MAR-2001; 2001US-0276822P.
 PR 18-MAR-2002; 2002US-00276822.
 XX
 XX (PANA-) PANACEA PHARM.
 PA
 XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compdre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabjohn PA, Shin DS, Stanley JS;
 XX
 DR WPI; 2003-018765/01.
 XX
 XX New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 PT
 XX Example 5; Page; 300pp; English.
 XX
 XX The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a food
 CC allergen, mutated to alter its IgE binding characteristics. Note: The
 CC present sequence is not shown in the specification but was created by the
 CC indexer using information provided in the specification
 XX
 SQ Sequence 626 AA;

 Alignment Scores:
 Pred. No.: 5,25e-304 Length: 626
 Score: 3280.00 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.08% Indels: 0
 DB: 6 Gaps: 0

 US-10-728-323-1 (1-2032) x ABUS2573 (1-626)

 QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGTGCTAGGATCCTGTCTGCTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20

 QY 110 TCTGCAACCGCATGCCAAGTCATCCTTACCAAGAGAAACAGAGAACCCCTGCCCCAG 169
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

 QY 170 AGGTGCTCCAGATTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCCAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60

 QY 230 CGCTGACCAAGCTCGAGTATGATCTCGTTGTGCTATGATCTCTCGAGGACACACTGGC 289
 DB

Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTTGGCCGCAACCCCGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGCGCATGGGGACCACTGGACCG 409
 Db 101 AspAspAspArgArgGlnProArgAlaGluGluGlyArgGlyProAlaGlyPro 120
 QY 410 AGGAGCGGTGAAGAAGAGAGACTGGAGACAACCAAGAGAGATTGGAGGCGCAAGT 469
 Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGCAGCCACGAAATAAGCCCGAAGAGAGAGAGAGAGACAGAGTGGGGAACA 529
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAAACATCTCGAAACAACCCCTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTAGACCCGCTACGGGAACCAAAACGGTAGATCCGGGTCTCTGCAGAGGTTTAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGCGTCAAGGCAGCTTTCAGAACTCCAGAACTCCAGAACTCCAGTATTGTGCAGATCGAGGCCAAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CTTAACTCTTGTCTTCTCCCAAGCAGCTGATGCTGATGATAACATCTTGTATTATCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACGCTGACCGTGAATAATGCGCAATAACAGAAAGAGGCTTTAACTTTCAGCAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCATGCACTCAGAAATCCCATCCGTTTCTATCTTCTACATCTTGAACCGCATCACAAC 889
 Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACACCCGCGCAGTTTCAGGAT 949
 Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTTCTCCCGGAGCAGCGGAGACCAATCATCTTCTACTTGCAGGGCTTCAGCGGAATACG 1009
 Db 301 PhePheProAlaSerSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGAGGCGCGCTTCAATGCGGAATTCAATGAGATACGAGGGTGCTGTTAGAAGAGAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGGCGAGGCGATCGAGTACTCTCGAGTAGTCAGAAC 1129
 Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGAGTGTATGATCAAGTGTCAAAGAGACGCTTGAAGAACTTACTTAAGCACCGT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGCTCAAGAAAAGGCTCCGAAGAAGAGGAGATATACCAACCAACCAATCACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAAGGCGAGCCCGATCTTTCTTAACAACTTTGGAAAGTTATTGTAGGTGAAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGACATGATGCTCAGCTGTGTAGAGATCAAGAA 1369
 Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCGTCTCGTCAACAAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAACTCGAAACCTTGAACCTGCTGCTGTAAGAAAGAGCAACAAAGAGGGGACGGCGG 1489
 DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnA-GlyArgArg 480
 QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAAGAGAGAGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
 DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValAlaArgArgTyr 500
 QY 1550 ACAGCGAGGTGAAGGAGGCGATGCTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATC 1609
 DB 501 ThrAlaAlaGlyLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCCAGAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCTTCAGGTGATAAGGACAAATGTGTATGACCATAGACAGATAGAGAACGAAGGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGTGGTGAACAGTTGAGAGCTCATCAAAACCAAGAGGATCT 1789
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTTGAGTGTCTGCTCTCAATCTCAATCTCGTCTCGTCTCTCTGAGAAGAG 1849
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTCAGAAAGAGATCAAGAGGAGGAAACCAAGAGGGAGGGTCCACTCTTTCA 1909
 DB 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTGAAGGCTTTTAAAC 1927
 DB 621 IleLeuLysAlaPheAsn 626
 RESULT 11
 ADM12053
 ID ADM12053 standard; protein; 626 AA.
 AC
 XX ADM12053;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Arachis hypogaea 1 (Ara hi) protein.
 XX
 KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
 KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
 KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
 KW antiinflammatory; antidiabetic; antithyroid; immune;
 KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
 KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
 KW scleroderma; dermatomyositis; pemphigus.
 XX
 OS Arachis hypogaea.
 XX
 XX US2003235594-A1.
 PN
 XX 25-DEC-2003.
 PD
 XX 17-SEP-2002; 2002US-00245871.
 XX
 XX 14-SEP-1999; 99US-00396813.
 PR
 XX 17-JUL-2002; 2002US-00197000.
 XX
 XX (ANTI-) ANTIGEN EXPRESS INC.
 PA
 XX Humphreys R, Xu M;
 PI

XX WPI; 2004-070554/07.
 DR
 XX Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
 PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
 PT erythematosus and diabetes mellitus.
 XX
 PS Example 1; Page 16-17; 87pp; English.
 XX
 CC The invention relates to a novel antigen presentation enhancing hybrid
 CC polypeptide. The novel polypeptide has an N-terminal element consisting
 CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
 CC deletion modifications, a chemical structure covalently linking the N-
 CC terminal element to an MHC class II-presented epitope of a C-terminal
 CC element. The C-terminal element comprises an antigenic epitope, which
 CC binds to an antigenic peptide binding site of an MHC class II molecule.
 CC The antigen presentation enhancing hybrid polypeptide has the following
 CC activities: antibacterial, virucide, fungicide, antirheumatic,
 CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
 CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
 CC enhancing hybrid polypeptide is useful for modulating the immune response
 CC in an individual and for treating infections (such as bacteria, virus,
 CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
 CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 CC represents a mammalian Ii key related protein of the invention.
 XX
 SQ Sequence 626 AA;

Alignment Scores:
 Pred. No.: 6,54e-304 Length: 626
 Score: 3279.00 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.06% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-1 (1-2032) x ADM12053 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTGCTAGGATCCTTCTCTGGTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACCGATGCCAAGTCATCCTTACCAGAGAAAACAGAGAACCCCTCGGCCAG 169
 DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCTCCAGAGTTGTCAACAGGAACCGGATGACTTCAAGCAAAAGGATCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlyLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTCGAGGACACACTGGC 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCTCCCTCAGGGGAGCGGACACGTGGCGCCCAACCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGCTCAACCCCAAGAGAGAGGAGGCGCGGACCGAGCTGGACCG 409
 DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGGTGAAGAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGGCGCAAGT 469
 DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
 QY 470 CATCAGCAGCCACGGAATAAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTTACTTCCGTCAGG 589

Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrrPheProSerArg 180
 QY 590 CGGTTTAGCACCCGCTACGGGAACCAACCGTAGGATCCGGTCTCTCGAGAGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAGGCGAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCCAGGCCAAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTGTGTTCTCCAGACGCTGATGCTGATAACATCTTGTGTTATCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTAGCCGTAGCAAAATGCCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCATTCAGAAATCCCATCCGGTTCATTTCCTACATCTTGAAACCCCATGACAAAC 889
 Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGGTTAACACACCCCGCCAGTTTGAGGAT 949
 Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCCTCCCGCGAGCGAGACCAATCATCTCTCTTCAGGGCTTCAGCAGGAAATACG 1009
 Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATGAGATACGGAGGTCGTGTAGAGAGAAAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
 Db 341 AlaGlyGlyGlnGlnGluArgGlyGlnArgTyrPheSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGGAGTGTAGTCAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 380
 QY 1190 AAATCCGTCTCAAAGAAAGCTCCGAAAGAGGGAGATATCACCAACCCCAATCAACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAAGAGCGAGCCCATCTTCTAACACTTTGGGAAGTTATTCAGGTGAAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AGAAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
 Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
 QY 1370 GGAGCTTTGATGCTCCCACTTCAACTCAAAGGCCATGTTATCGTCTGTCACAAA 1429
 Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
 QY 1430 GGNACTGGAAACCTTGAACCTGGCTGTAGAAGAGCAACACAGAGGGCGCGCG 1489
 Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
 QY 1490 GAAAGAGAGGAGGACCAAGACCAAGAGAGGAGGAAAGTAAACAGAGAGTGGCTAGGTAC 1549
 Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGGTGAAGAGCGGATGTGTTTCATCATGCCAGAGCTCATCCAGTAGCCATC 1609
 Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AAGCTTCTCCGAATCTCATCTGCTTGGCTTCGGTATCAAGCTGAAACACACAGAGA 1669
 Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540

QY 1670 ATCTTCTTGCAGGTGATAAGGACAAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGAT 1729
 Db 541 IlePheLeuAlaGlyValPlyAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTGCGGTGAAACAAGTTGAGAAGCTCATCAAAAACAGAAAGAAATCT 1789
 Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTGTGAGTGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAG 1849
 Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGlnLysGlu 600
 QY 1850 TCTCTCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
 Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927
 Db 621 IleLeuLysAlaPheAsn 626
 RESULT 12
 ABUS2569
 ID ABUS2569 standard; protein; 626 AA.
 XX AC ABUS2569;
 XX XX
 DT 10-MAR-2003 (first entry)
 XX XX
 DE Peanut Ara h1 mutant D52A.
 XX XX
 KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
 KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
 KW wound healing.
 XX XX
 OS Homo sapiens.
 XX XX
 PN WO200274250-A2.
 XX XX
 PD 26-SEP-2002.
 XX XX
 PF 18-MAR-2002; 2002WO-US009108.
 XX XX
 PR 16-MAR-2001; 2001US-0276822P.
 PR 18-MAR-2002; 2002US-00276822.
 XX XX
 PA (PANA-) PANACEA PHARM.
 XX XX
 PI Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;
 PI Compadre CW, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabjohn PA, Shin DS, Stanley JS;
 XX XX
 DR WPI; 2003-018765/01.
 XX XX
 PT New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 XX XX
 PS Example 5; Page; 300pp; English.
 XX XX
 CC The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,

CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification.
XX
SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 8.15e-304 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 92.03% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABUS2569 (1-626)

QY 50 ATGAGAGGAGGGTTCTCCACGTATGCTGTCTAGGATCCTTGTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATCCCAAGTCATCACCTTACCAGAGAAACACAGAACCCCTGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysThrGlnAsnProCysAlaGln 40
QY 170 AGTGCTCTCCAGATGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAlaLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGCAACAGTCAGATGATGATCTCTGTTGTCTATGATCCTCGAGACACATGCG 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGACACAGTGGCGCCCAACCCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGAGGCGGATGGGGACCATCGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGACGCTGAAGAGAGAGAGTGGAGACCAACAGAGAGATGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
QY 470 CATCAGCAGCCAGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGlnGlnTyrGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGACCCGCTACGGAAACCAAAACCGTAGGATCCGGCTCTCGAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGCAGTTTCAGATCTCCAGATCTCCAGATCTCCAGATCTCCAGATCT 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCAAAGCAGCTGATGCTGATAAATCTTGTATCTCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCACAGCCCGTGAACCGTAGCAATGCGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCCATCCCGTTTCATCTTCTACATCTTGACCGCCATGACAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCCGTAAACACACCCCGCCAGTTTGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCGGAGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCCCTTCAATGCGGAATTCATAGATACGAGGGTGTCTGTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGTGAACAAG 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgGlyGlnArgGlyGlnArgGlyGlnArg 360
QY 1130 AATGAGGAGTATAGTCAAGTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCAAGAAAGAGGCTCCGAAGAGAGAGAGAGATATCACAACCCCAATCACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGGCGAGCCCGCATCTTTCTAACTTTGGGAAGTATTTTGTAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGCTTCAGGACCTGGACATGATGCTCACCCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATCTGCTCGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValLys 460
QY 1430 GGAATCGGAACCTTGAATCTGCTGTGTAAGAAAGAGCAACAACAGAGGGAGCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAG 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGAAAGCGCATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCATCTGCTTGGCTTCCGATCAACGCTGAAACCAACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCCAGGTGATAAGGACATGTGTAGACCAGATAGAGAGAGAGAGAGAGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTCCGGTGAACAGTTGAGAAGCTCATCAAAACACAGAAAGAACT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCGTCCCTCAATCTCAATCTCGTCTCGTCTCGTCTCGAGAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927
DB 621 IleLeuLysAlaPheAsn 626

[illegible]

Db 361 AsnGluGlyValIleValIleValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCCTCTCAAGAAGGCTCCGAGAAGAGGAGATATACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGGCGGAGCCGATCTTTCTACAACCTTTGGGAAGTTATTGAGGTGAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValIleYsProAsp 420
QY 1310 AAGAAGAACCCAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTCACTCAAGCCCATGTTATGTCGTCGCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGCGACGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
QY 1490 GAAGAAGAGGAGGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgGly 500
QY 1550 ACAGCAGGTTGAAGGAGCGGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCCTCCGACCTCCATCTGCTGGCTTCGCTATCAACGCTGAAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTGACGTCATAGGACATGTGTATGACACAGATAGACGAGCAAGCAAGGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGAGTGAACAGTTGAGAGCTCATCAAAACACAGAGGAGATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 14
AAW22150
ID AAW22150 standard; protein; 626 AA.
XX
AC AAW22150;
XX
DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hi.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hi.
XX
OS Arachis hypogaea; strain Florunner.
XX
FH Key Location/Qualifiers
FT Peptide 1..22

FT Protein /label= Sig_peptide
FT 23..626
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX
PN W097241139-Al.
XX
PD 10-JUL-1997.
XX
XX 23-SEP-1996; 96WO-US015222.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 04-MAR-1996; 96US-00610424.
XX
PA (UYAR-) UNIV ARKANSAS.
PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
XX WPI; 1997-363453/33.
DR N-PSDB; AAT76613.
XX
PT Peanut allergens Ara hi and Ara hII - used for vaccination and in two-
PT site monoclonal antibody based ELISA.
XX
PS Claim 31; Page 172; 354pp; English.
XX
CC This polypeptide comprises major peanut allergen Ara hi (AAW22149). Its
CC sequence was deduced from cDNA clone P41b (AAT76613), isolated from
CC peanut seed cDNA using a primer (see AAT76616) based on an isolated Ara
CC hi peptide (see AAW24206). The sequence shows significant homology with
CC the vicilin family of seed storage proteins of other legumes. The
CC allergen is recognised by serum IgE from a large proportion of
CC individuals with peanut hypersensitivity. Ara hi and Ara hII (see
CC AAW24164) can be used to raise monoclonal antibodies which are used in a
CC specific two-site MAb ELISA for the detection of Ara hi or Ara hII
CC (claimed). IgE-binding Ara hi antigen epitopes (see AAW24165-87) may be
CC used in vaccines to protect against allergic reactions to peanut
CC allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to
CC standardise OS field)
SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 1.97e-303 Length: 626
Score: 3274.00 Matches: 624
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 2
Query Match: 91.91% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x AAW22150 (1-626)

QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTAGGATCCTTGTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCCAAGTCATCATCTTACAGAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCTCAGAGTTGTCAACAGGACCGGATAGCTTGAAGCAAAAGCATGCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTCCACCAAGCTCGAGTATGATCCTCGTGTCTGTATGATCCTCGAGGACACACTGCG 289
Db 61 ArgCysThrLysLeuLeuLutyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTTCCCTCCAGGGAGCGACACAGTGGCGCGCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlnProGlyAspTyr 100

QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGAGCGCGATGGGGACCACTGGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGACGCTGAAGAGAACGACTGGACACCAACCAAGAGAGATTGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCACCCCGAAGAAATAAGGCGCCGAAGAGAGAGAGAAACAAGATGGGGGAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CGAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACCGTAGGATCGGGTCTCTGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATTCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACGCTGACCGTAGCAATGCGCAATAACAGAAAGAGCTTTAATCTTACAGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCATCCGGTTCATTTCTACATCTTGAAACCGCCATGACAAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAACACACCGCGCCAGTTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValThrGlnProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCGAGACCAATCATCTACTTGCAGGGCTTCAGAGGAAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCGCTTCAATCCGAATTTCAATGAGATACGGAGGCTGTGTTAGAAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGTGAGCAAGAGAGAGAGCGGACGCGATGAGTACTCGGAGTGTAGTAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTGTATGATCAAGTGTCAAGGAGGACGTTGAAGACTTACTTAAGCACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGCGGACCGCATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAACCCCGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 CGAGCTTTGATGCTCCACACTTCAACTCAAAGGCGCATGTTTATCGTCTGCTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGG 1489

DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 CAAAGAAGAGGAGGACGAAGACGAAAGAGAGGAGGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTTCAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTTCTCGAACTCCATCTGCTTGGCTTCGTATCAAGCTGAAACACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGCAACATGTGTAGACCAGATAGACAAGCAAGCGAAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysLeuIleLysGlnLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCTCTGAGAAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAGAGGATCAAGAGGAGGAAAACCAAGGAGGAGGGTCCACTCTCTTCA 1909
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
DB 621 IleLeuLysAlaPheAsn 626
RESULT 15
ADG27464
ID ADG27464 standard; protein; 625 AA.
XX
AC ADG27464;
XX
DT 26-FEB-2004 (first entry)
DE Peanut Ara h1 cDNA clone P41b protein.
XX
KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
KW antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.
XX
OS Arachis hypogaea.
XX
PN US2003202980-A1.
XX
PD 30-OCT-2003.
XX
PF 18-MAR-2002; 2002US-00100303.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.
PR 13-FEB-1998; 98US-0074633P.
PR 29-JUN-1998; 98US-00106872.
PR 27-AUG-1998; 98US-00141220.
PR 13-NOV-1998; 98US-00191593.
PR 29-JAN-1999; 99US-00240557.
PR 29-JAN-1999; 99US-00241101.
PR 11-FEB-1999; 99US-00248673.
PR 02-MAR-1999; 99US-00248674.
PR 02-MAR-1999; 99US-0122450P.
PR 02-MAR-1999; 99US-0122452P.
PR 02-MAR-1999; 99US-0122560P.
PR 02-MAR-1999; 99US-0122565P.

Db 361 AsnGluGlyValIleValIysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAGAAAGCGTCCGAAAGAGGAGATATCACCAACCCCAATCAACTTG 1249
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGCGAGCCCGATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AGAGAAGACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGATCAAGAA 1369
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTTCATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCACAAA 1429
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAGAAAGAGCAACAACAGAGGGGACGGCGG 1489
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGAGGACCAAGACAGAGAGAGAGGGAGTAACAGAGAGGTGCGTAGGTAC 1549
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGGAAGCGGATGTGTTTCATCGCCAGCAGCTCATCCAGTAGCCATC 1609
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAACAGA 1669
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTTGCAGCTCATPAGGACATGTGTAGACCAGATAGAGAGCAACCGAAGGAT 1729
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 IlePheLeuAlaGly---LysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 559
QY 1730 TTAGCATTCCTCGGTGGGTGAACAAGTTGAGAAGCTCATCAAAACCCAGAAAGGAATCT 1789
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
560 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 579
QY 1790 CACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAG 1849
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
580 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 599
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
600 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 619
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
DB |||||||||||||||||||
620 IleLeuLysAlaPheAsn 625

Search completed: August 24, 2005, 09:55:54
Job time : 154.082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 07:15:51 ; Search time 25.4416 Seconds
(without alignments)
15369.504 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3562

Sequence: 1 aataacatcatatcatc.....cggtgtgcgtgtttctcc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPO.spool/US10728323/runat_23082005_124355_29222/app_query.fasta_1.4757
-DB=PIR_79 -QPMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323@cgn_1_1_76 @runat_23082005_124355_29222 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	36.0	605	2 S20007	beta-conglycinin a
2	1281	36.0	605	1 FMSYBA	beta-conglycinin a
3	1267	35.6	571	2 S00566	convicillin precurs
4	1251	35.1	639	2 B24810	beta-conglycinin a
5	1178.5	33.1	459	2 S08505	vicilin - garden p
6	1174.5	33.0	433	2 S00567	vicilin precursor
7	1156	32.5	439	1 FMSYCB	beta-conglycinin b
8	1148	32.2	438	2 S35757	vicilin, 4YK - gar
9	1109.5	31.1	463	2 A27288	vicilin precursor
10	1108.5	31.1	463	2 S06309	vicilin precursor
11	1068	30.0	410	1 FPMNVB	vicilin B precursor
12	962.5	27.0	445	2 JQ2264	canavalin - jack b
13	959.5	26.9	445	2 S00281	canavalin - sword
14	907.5	25.5	810	2 T44430	protein PV100 [imp

15	839	23.6	605	2 S06398	alpha-globulin typ
16	825.5	23.2	436	2 A23498	phaseolin type alp
17	809	22.7	566	2 S22477	vicilin precursor
18	806.5	22.6	430	2 S10156	alpha-phaseolin pr
19	788	22.1	421	2 A24810	phaseolin beta cha
20	786	22.1	509	2 S08059	alpha-globulin typ
21	778	21.8	588	1 FWCNAB	alpha-globulin B p
22	728.5	20.5	275	1 FPMVA	vicilin A precursor
23	629.5	17.7	582	2 B53234	vicilin-like stora
24	629	17.7	573	2 AS3234	globulin-18, GUB18
25	614	17.2	637	2 S35221	globulin Begi prec
26	587	16.5	218	2 S16334	beta-conglycinin a
27	569	16.0	450	2 S15675	globulin-2 precurs
28	569	16.0	524	2 JQ1730	62K sucrose-bindin
29	562	15.8	483	2 T06459	62K sucrose-bindin
30	559.5	15.7	448	1 S24756	vicilin-like stora
31	536.5	15.1	240	2 T06399	7S storage protein
32	470.5	13.2	196	2 S02281	convicillin (clone
33	463.5	13.0	540	2 S21825	vicilin-like stora
34	403	11.3	414	2 S60190	vicilin - Zamia fu
35	397.5	11.2	296	2 T06572	convicillin precurs
36	336	9.4	184	1 FSPFB	phaseolin - kidney
37	291	8.2	499	2 T10443	probable major pro
38	284	8.0	124	1 FPMVA	vicilin, 14K compo
39	284	8.0	407	2 T02258	globulin1 - maize
40	269.5	7.6	236	2 T01662	globulin1 - maize
41	269.5	7.6	699	2 E84565	hypothetical prote
42	263.5	7.4	511	2 E84685	probable seed stor
43	260.5	7.3	481	2 T14300	hypothetical prote
44	228.5	6.4	499	2 D85433	globulin-like prot
45	211.5	5.9	516	1 FWCNBB	beta-globulin B pr

ALIGNMENTS

RESULT 1

S20007 beta-conglycinin alpha chain precursor - soybean

C:Species: Glycine max (soybean)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S20007

R:Lelievre, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.

Plant Mol. Biol. 18, 259-274, 1992

A:Title: Synthesis and assembly of soybean beta-conglycinin in vitro.

A:Reference number: S20007; M0ID:92119248; PMID:1731988

A:Accession: S20007

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-605 <LE>

A:Cross-references: UNIPROT:Q94LX2

C:Superfamily: glycinin

Alignment Scores:

Pred. No.:	8.03e-82	Length:	605
Score:	1282.00	Matches:	285
Percent Similarity:	59.3%	Conservative:	108
Best Local Similarity:	43.05%	Mismatches:	173
Query Match:	35.99%	Indels:	96
DB:	2	Gaps:	18

US-10-728-323-1 (1-2032) x S20007 (1-605)

QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGATCTTGTCTGCTTCAGTT	109
DB	2	MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
QY	110	TCTGCAACGATGCCATCATCTTACACAGAGAAACAGAACCCCTGCCGCCAG	169
DB	20	SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn	36
QY	170	AGGTGCTCCAGTGTGTCAACAGGACCGGATGACTTGAACAAAGCATCGCGAGTCT	229
DB	37	LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56

```
QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
D  |||||
D  57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluClyGluileProArg--- 75
QY 263 GTCTATGATCTCGAGGACACACTGGCACCACCAACCAAGTTCCCTCAGGGAGCGG 322
D  |||||
D  76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACACGTGGCGCCCAACCCGAGACTACGATGAC----- 358
D  |||||
D  85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluGlnProArgProGln 104
QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAAGAGAGCCGA 391
D  |||||
D  105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCAAGCTGGACCGAGGAGCGTGAAGAGAGAGAA 430
D  |||||
D  125 TrpProArgLysGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144
QY 431 GACTGGAGACAACCAAGAGAGAT-----TCGAGGCGCAAGTCAATCAGCAGCCACGG 484
D  |||||
D  145 AspGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu 164
QY 485 AAAATAAGGCCCGAAGAGAGAGAGAACAGAGTGGGGAAACACACAGGTAGCCATGTG 544
D  |||||
D  165 ArgLysGlnGluGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGAA-----GAAACATCTCGGAACAACCCCTTCTACTTCCGTCACAGGCGTTTACG 598
D  |||||
D  185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCCGTACGGGAACCAACAGTAGATCCGGTCTCGAGGTTTCGACCAAGGTCA 658
D  |||||
D  205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGGCAGTTTCAGAAATCTCCAGATACCCGTATTGTGCAGATCGAGCGAAACCTAACACT 718
D  |||||
D  225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTGTGTTCTCCCAAGCAGCTGATGATACATCTTGTATTATCAGCAAGGCGAAGCC 778
D  |||||
D  245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAGGSCCATGCA 838
D  |||||
D  265 IleUserLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCATGACAAACAGAACCTC 898
D  |||||
D  285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
D  |||||
D  305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY 959 GCGAGCAGCGGACCAATCATCTACTTGCAGGGCTTCAGAGGAATACGTTGAGGCGC 1018
D  |||||
D  325 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAAATTCATAGATACGAGGAGTGTCTGTAGAAGAAATCGAGAGGT 1078
D  |||||
D  345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGly--- 363
QY 1079 GAGCAAGAGAGAGAGGCGAGGCGCATGGAGTACTCGGAGTAGTGAGAACCAATGAAGA 1138
D  |||||
D  364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCACCGCTAAATCCGTC 1198
D  |||||
D  375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
```

```
QY 1199 TCAAAAGAAAGCGTCCGAAGAAGAGGAGATATCATCAACCAACCAATCAACTTGAGAGAAGC 1258
D  |||||
D  395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTCTAAACAACCTTGGGAAGTATTATTGAGTGAACGACAGAGAGAAC 1318
D  |||||
D  412 AspProIleTyrSerAsnLysLeuGlyLysPheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGAGCTCGACATGATGCTCACTCTGTAGAGATCAAAAGAGAGGAGCTTTG 1378
D  |||||
D  431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCACACTTCAACTCAAAAGCCATGTTATCGTCGTGTCAACAAAGAACTGGA 1438
D  |||||
D  451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCGTGGCTGTAAGAAAGAGCACAACAGAGGGGACGGCGGGAAGAAGAG 1498
D  |||||
D  471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGluGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAAGAGGAGGAAATTAACAGAGAGGTGCGTAGGTACACAGCGAGG 1558
D  |||||
D  490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAAGCGATGTGTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
D  |||||
D  500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValAsnAlaThr 519
QY 1619 TCCGAACCTCCATCTCTGCTTGGCTTGGTATCAACGCTGAAACCAACACAGATCTTCCTT 1678
D  |||||
D  520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGTATAGGACAATGTGTATAGACACCATAGAGAAGCAAGCAAGATTTAGCATTC 1738
D  |||||
D  540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTGGCTCGGTGAAACAAGTTTCAGAAAGCTCATCAAAAAACCAAGAGGAATCTCACTTTGTG 1798
D  |||||
D  560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGCTCTCTCAATCTCAATCTCAATCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAG 1858
D  |||||
D  580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGCTCCACTCTCTTCAATTTTGAAG 1918
D  |||||
D  586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
D  |||||
D  603 AlaPhe 604
RESULT 2
FWSYBA
beta-conglycinin alpha chain precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: S14681; S74124; S06714
R:Sebastiani, F.L.; Farrell, L.B.; Schuler, M.A.; Beachy, R.N.
Plant Mol. Biol. 15, 197-201, 1990
A:Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin.
A:Reference number: S14681; MUID:91355860; PMID:2103438
A:Accession: S14681
A:Molecule type: mRNA
A:Residues: 1-605 <SEB>
A:Cross-references: UNIPROT:P13916; EMBL:X17698; NID:g18535; PIDN:CAA35691.1; PID:g18536
R:Shutov, A.D.; Kakhovskaya, I.A.; Bastrygina, A.S.; Bulmaga, V.P.; Horstmann, C.; Muent
Eur. J. Biochem. 241, 221-228, 1996
A:Title: Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage gl
A:Reference number: S74123; MUID:97054613; PMID:8898910
A:Accession: S74124
A:Molecule type: protein
A:Residues: 189-196,'H','198','N','200','X','202-203;397-408','X','410','X','412-417','X','419-420',
```

A: Experimental source: seed	
C: Superfamily: glycinin	
F: 1-22/Domain: signal sequence #status predicted <SIG>	
F: 23-62/Domain: propeptide #status predicted <PRO>	
F: 63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>	
F: 261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted	
Alignment Scores:	
Pred. No.:	9,43e-82
Score:	1281.00
Percent Similarity:	59.21%
Best Local Similarity:	43.05%
Query Match:	35.96%
DB:	1
US-10-728-323-1 (1-2032) x FWSYBA (1-605)	
QY	50 ATGAGAGGGAGGTTTCTCCACTGATGCTGCTGCTAGGATCCTGTGCTGGCTTCAGTT 109
DB	2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY	110 TCTGCAACGCGATCCCAAGTCACTCACCTTACCAGAAGAAACACAGAGAACCCCTCGGCCAG 169
DB	20 SerValSerPheGlyIleAla-----TyrTipGluLys---GluAsnProLysHisAsn 36
QY	170 AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGGATCT 229
DB	37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY	230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
DB	57 ArgCysAsnLeuLeuLysValGluLysGluLysGluLysGluLysGluLysProArg--- 75
QY	263 GTCTATGATCCTCGAGGACACACTGGCACCAACCAACCACTTCCCTCCAGGGAGCGG 322
DB	76 -----ProArgProArg-----ProGlnHisProGlu 84
QY	323 ACAGTGGCCGCCAACCGGAGACTACGATGATGAC----- 358
DB	85 ArgGluProGlnGlnProGlyGluLysGluLysGluLysGluLysGluLysProArgProIle 104
QY	359 -----CGCGTCAACCCCAAGAGAG-----GAAGAGGCCGA 391
DB	105 ProPheProArgProGlnProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 124
QY	392 TGG-----GGACCAGCTGGACCGGAGGCGGTGAAGAGAGAA 430
DB	125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
QY	431 GACTGAGACAACCAAGAGAGAT-----TGGAGCGACCAAGTCAATCAGCAGCCACGG 484
DB	145 AspGluGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
QY	485 AAAATAAGGCCCAAGGAGAGAGAGACAGAGTGGGACACACAGTACGATCGCATGTG 544
DB	165 ArgAsnGluGluGluAspGluAspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 184
QY	545 AGGGAA-----GAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGCGGTTTAGC 598
DB	185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY	599 ACCCGTACGGGAACCAACCGTAGATCCGGTCTCTGAGAGTTTGAACCAAGTCA 658
DB	205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY	659 AGCAGTTTCAGATCTCCAGATCACCGTATTGTGCAGATCCAGGCCAACCTACACT 718
DB	225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY	719 CTTGTTCTTCCCAAGCAGCTGATGATGATAAATCCTTGTATTCAGCAAGGCGCAAGCC 778
DB	245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264

QY	779 ACCGTACCGGTAGCAAAATGCGCAATTAACAGAAAGAGAGCTTTAAATCTTACGAGGGCCATGCA 838
DB	265 IleLeuSerLeuValAsnAsnAspAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY	839 CTCAGAAATCCCATCCGGTTTCTATCTTCTACATCTTGAACCGCCATCAGCAACAGAACCTC 898
DB	285 LeuArgValProSerGlyThrThrTyrValValAsnProAsnAsnGlnAsnLeu 304
QY	899 AGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCCGGCAGTTTTCAGGATTTCTTCCCG 958
DB	305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
QY	959 GCGAGCAGCGAGACCAATCATCTTCTTGCAGGGCTTCAGAGGAATAGTTTGGAGGCC 1018
DB	325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY	1019 GCCTTCAATCGGAATTCAATGAGATACGAGGGTCTCTTAGAGAGAGAAATGCAGGAGGT 1078
DB	345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGly--- 363
QY	1079 GAGCAAGAGGAGAGGCGAGCGCATGGAGTACTCTCGAGTAGTGAGAACAAATGAAGGA 1138
DB	364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
QY	1139 GTGATAGTCAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCACGCTAAATCCGTC 1198
DB	375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY	1199 TCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCACTTTGAGAGAGGC 1258
DB	395 SerArgLysThrLysSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY	1259 GAGCCCGATCTTTTAAACAACCTTGGGAGTATTATTGAGGTGAAGCCAGCAAGAGAAC 1318
DB	412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY	1319 CCCAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGATCAAAAGAGAGCTTTG 1378
DB	431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY	1379 ATGCTCCACACTCAACTCAAGGCCATGGTTATCGTCTGTCACCAAGGAACTGGA 1438
DB	451 LeuLeuProHisPheAsnSerLysAlaIleValIleValIleAsnGluGlyAspAla 470
QY	1439 AACCTTGAACCTCGTCTGTAAGAAAGAGCAACCAACAGAGGGGCGCGGGAAGAGAG 1498
DB	471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGln 489
QY	1499 GAGGAGAAAGCAAGAGAGAGGAGGAGTAACAGAGGTCGTTAGGTACACAGCGAGG 1558
DB	490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY	1559 TTGAAGGAGCGATGCTGTTTCATATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTC 1618
DB	500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY	1619 TCCGAACCTCCATCTGCTGGCTTCGTTATCAACGCTGAAACCAACCAAGAAATCTTCTCT 1678
DB	520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY	1679 GCAGGTGATAGCAATGTGTATAGACCCAGATAGAGACGAGCGGAGAGATTAGCATTC 1738
DB	540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGlnLeuAlaPhe 559
QY	1739 CTTGGTCTGGGTAACAAGTTTCAGAGGCTCATCAAAACCAAGCAAACTCTCCTTTG 1798
DB	560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGlnSerTyrPheVal 579
QY	1799 AGTGCTCGCTCAATCTCAATCTCGTCTCTCTCTCGAGAAAGAGTCTCTCTGAG 1858
DB	580 AspAlaGlnProLysLys----- 585

Db	397	GlulGlnGlnGlnGlnGluGluArgLysGln	407
Qy	1127	AACAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTTAAGCAC	1186
Db	408	-----GlnSerValIleValGluIleSerLysGlnIleArgGluLeuSerLysHis	425
Qy	1187	GCTAAATCCGTCTCAAAGAAAGGCTCCGAGAGAGAGGAGATATCAACCAACCCAATCAAC	1246
Db	426	AlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPheAsn	442
Qy	1247	TTGAGAAAGCGCAGCCGATCTTTCAACACTTTGGGAAGTTATTTGAGGTCAAGCCA	1306
Db	443	LeuLysSerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIle-----	460
Qy	1307	GACAAGAAAGACCCCAAGCTTCAGGACCTGGACATGATGTCACTCTGTAGAGCAAAA	1366
Db	461	ThrGlnArgAsnProGlnLeuArgAspLeuAspValPheLeuSerValValAspMetAsn	480
Qy	1367	GAGGAGCTTTGATGCTCCACACTTCAACTCAAAAGGCCATGGTTATTCGTGCGTCAAC	1426
Db	481	GlulGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIleAsn	500
Qy	1427	AAAGGAACTGGAACCTTGAACCTCGTCGCTGAAGAAAGCAACCAACAGAGGGACGG	1486
Db	501	GlulGlyGluAlaAsnIleGluLeuValGlyLe-----LysGluGlnGlnGlnArgGlnGln	519
Qy	1487	CGGAAAGAGAGGAGGACGAAGCAAGAGAGGAGGCGAAGTAACAGAGAGGTGCGGTAGG	1546
Db	520	GlnGluGlnPro-----LeuGluValArgLys	529
Qy	1547	TACACAGCAGGTTGAAGGAAGGCGATGTGTTCAATCAATGACGAGCTCATCAGTAGCC	1606
Db	530	TyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValMet	549
Qy	1607	ATCAACGCTTCCCGAACTCCATCTCTGCTTGGCTTCGGTATCAACGCTGAAAACACCAC	1666
Db	550	ValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsnGln	569
Qy	1667	AGAATCTTCCTTCGAGGTGAAGGACAAATGCTGATACACGATAGAGAAGCAAGCAAG	1726
Db	570	ArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnValGln	589
Qy	1727	GATTTAGCATTCCTGGGTGGGTGAACAAGTTGAGAAGCTCATCAAAAAACGAGAGGAA	1786
Db	590	GluLeuAlaPheProArgSerAlaLysAspIleGluAsnLeuIleLysSerGlnSerGlu	609
Qy	1787	TCTCACPTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCTGAGAAA	1846
Db	610	SerTyrPheValAspAlaGlnProGln-----	618
Qy	1847	GAGTCTCTGAGAAAGGAGTCAAGGAGGAGGAAAAACCAAGGAGGAGGAGGTCCACTCTCTT	1906
Db	619	-----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeuSer	632
Qy	1907	TCAATTTTGAAGCGTTTTT	1924
Db	633	SerIleLeuArgAlaPhe	638

RESULT 5
S08505
vicilin - garden pea
C;Species: Pisum sativum (garden pea)
C;date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: S08505
R;Higgins, T.J.V.; Newbigin, E.J.; Spencer, D.; Llewellyn, D.J.; Craig, S.
Plant Mol. Biol. 11, 683-695, 1988
A;Title: The sequence of a pea vicilin gene and its expression in transgenic tobacco plants
A;Reference number: S08505
A;Accession: S08505
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <HIG>

A;Cross-references: UNIPROT: P13918; EMBL: X14076; NID: g20917; PIDN: CAA32239.1; PID: g31367
C;Genetics:
A;Introns: 102/1; 160/3; 187/3; 295/3; 400/1
C;Superfamily: glycinin

Alignment Scores:
Pred. No.: 1.33e-74 Length: 459
Score: 1178.50 Matches: 223
Percent Similarity: 71.33% Conservative: 93
Best Local Similarity: 50.34% Mismatches: 114
Query Match: 33.09% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S08505 (1-459)

Qy	542	GTGAGGGAAGAAACATCTCGGAACACCTTTTCTACTTCCGCTCAAGGCGTTTAGCACC	601
Db	26	ValSerSerArgSerAspProGlnAsnProPheIlePheLysSerAsnLysPheGlnThr	45
Qy	602	CGTACGGGAACCAAAACGGTAGGATCCGGGTCCTCAGAGGTTTGACCAAGGTCAAGG	661
Db	46	LeuPheGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnArgSerLys	65
Qy	662	CAGTTTCAGATCTCCAGAAATCACCGTATGTGCAGATCGAGGCCAAACCTTAACACTT	721
Db	66	IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIle	85
Qy	722	GTTCTTCCCAAGCAGCTGATGCTCATACATCCTTGTATTCCAGCAAGGGCAAGCCACC	781
Db	86	PheLeuProGlnHisThrAspAlaAspTyrIleLeuValValLeuSerGlyLysAlaIle	105
Qy	782	GTGACCGTAGCAATTCGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCAC	841
Db	106	LeuThrValLeuLysProAspAspArgAsnSerPheAsnLeuGluArgGlyAspThrIle	125
Qy	842	AGAATCCCATCCGTTTCATTTCTTACATCTTTGAACCGCCATGACCAACCGACCTCAG	901
Db	126	LysLeuProAlaGlyThrIleAlaTyrLeuValAsnArgAspAsnGluGluLeuArg	145
Qy	902	GTAGCTAAATCTCCATCCGCTTAAACACACCGCGCCAGTTTGAGGATTTCTCCCGCG	961
Db	146	ValLeuAspLeuAlaIleProValAsnArgProGlyGlnLeuGlnSerPheLeuLeuSer	165
Qy	962	AGCAGCCGAGACCAATCATCTTACTTTCAGGCGCTTCAGCAGGAATACGTTGAGGCGCC	1021
Db	166	GlyAsnGlnAsnGlnAsnTyrLeuSerGlyPheSerLysAsnIleLeuGluAlaSer	185
Qy	1022	TTCAATGCGGAATTCATAGATACGAGGCGTCTGTAGAACAGAAATCGCAGAGGTGAG	1081
Db	186	PheAsnThrAspTyrGluGluIleGluLysValLeuLeuGluGluHisGluLysGluThr	205
Qy	1082	CAAGAGNAGAGGCGCAGCGGATCGGAGTACTCGGAGTAGTCAGAACAAATGAAGAGTG	1141
Db	206	GlnHisArgSerLeuLys-----AspLysArgGlnGlnSerGlnGluGluAsnVal	223
Qy	1142	ATAGTCAAAGTGTCAAAGCAGCAGCTTCAAGAACTTACTAAGCAGCTAAATCCGCTCA	1201
Db	224	IleValLysLeuSerArgGlyGlnIleGluLeuSerLysAsnAlaLysSerThrSer	243
Qy	1202	AAGAAAGGCTCCGAAGAAGAGGAGATATCAACACCCCAATCAACTTGTAGAGAGGCGAG	1261
Db	244	LysLysSerValSerSerGluSerGlu-----ProPheAsnLeuArgSerArgGly	260
Qy	1262	CCCGATCTTTTACAACTTTGGGAGTTATTGTAGGTCAAGCCGACACAGAGACCCC	1321
Db	261	ProTleTyrSerAsnGluPheGlyLysPhePheGluIleThrProGlu---LysAsnPro	279
Qy	1322	CAGCTTCAGGACCTCGGACATCATGTCTACCTCGTAGAGATCAAGAAAGAGGAGCTTGTATG	1381
Db	280	GlnLeuGlnAspLeuAspIlePheValAsnSerValGluIleLysGluGlySerLeuLeu	299
Qy	1382	CTCCCACTTCAACTCAAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGAAAC	1441

Db 300 LeuProHisTyrAenSerArgAlaIleValThrValaenGluGlyLysGlyAsp 319
QY 1442 CTTGAACCTCTGGCTGTAAAGAAAGACCAACAGAGGGAGCGCGGGAAGAGGAG 1501
Db 320 PheGluLeuValGlyGlnArgAenGluAenGlnGlnGlnArgLysGluAspGlu 339
QY 1502 GACCAAGACCAAGAGAGGAGGAGGAGTAACAGAGAGGTGGTAGTACACAGCAGGTTG 1561
Db 340 GluGluGlnGlnGlyGluGluGluIleAenLysGlnValGlnAenTyrLysAlaLysLeu 359
QY 1562 AAGGAAGCGGATGTTTCATCATGCCAGACGCTCATCCAGTAGCCATCAACGCTTCCTCC 1621
Db 360 SerSerGlyAspValPheValIleProAlaGlyHisProValAlaValLysAlaSerSer 379
QY 1622 GAATCCATCTGCTGGCTCGGTATCAACGCTGAAACCAACACAGATCTTCCTTGCA 1681
Db 380 AsnLeuAspLeuGlyPheGlyIleAsnAlaGluAsnGlnAsnGlnPheLeuAla 399
QY 1682 GGTGATAGACCAATGTAGTACACAGATAGAGCAAGCAAGGATTTAGCATTCCT 1741
Db 400 GlyAspGluAenValIleSerGlnIleGlnArgProValLysGluLeuAlaPhePro 419
QY 1742 GGTCGGGTGAACAGTTGAGAGCTCATCAAAACCAAGAGAAATCTCACTTTGTGAGT 1801
Db 420 GlySerAlaGlnGluValAspArgIleLeuGluAenGlnLysGlnSerHisPheAlaAsp 439
QY 1802 GCTCGTCTCAATCTCAATCTCAATCTCGTCTGCTCGTCTGAGAAAGTCTCTTGAGAA 1861
Db 440 AlaGlnProGlnGlnArg-----GluArgGlySerArgGluThr 452
QY 1862 GAGGATCAA 1870
Db 453 ArgAspArg 455

RESULT 6
S00567
vicilin precursor (clone pDUB9) - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S00567
R:Watson, M.D.; Lambert, N.; Delauney, A.; Yarwood, J.N.; Croy, R.R.D.; Gatehouse, J.A.;
Biochem. J. 251, 857-864, 1988
A:Title: Isolation and expression of a pea vicilin cDNA in the yeast Saccharomyces cerevisiae
A:Reference number: S00567; MUID:88326226; PMID:3046604
A:Accession: S00567
A:Molecule type: mRNA
A:Residues: 1-433 <WAT>
A:Cross-references: UNIPROT:P13918; EMBL:Y00722; NID:g20915; PIDN:CAA68708.1; PID:g75824
C:Superfamily: glycinin
F:1-2/Domain: signal sequence (fragment) #status predicted <SIG>
F:3-433/Product: vicilin #status predicted <WAT>

Alignment Scores:
Pred. No.: 2-5e-74 Length: 433
Score: 1174.50 Matches: 221
Percent Similarity: 72.18% Conservative: 93
Best Local Similarity: 50.80% Mismatches: 108
Query Match: 32.97% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S00567 (1-433)

QY 566 AACCTTTCTACTCCGCTCAAGCGGCTTTAGCACCGCTACCGGGAACCAACCGTAGG 625
Db 8 AsnProPheIleLysSerAsnLysPheGlnThrLeuPheGluAenGluAenGlyHis 27
QY 626 ATCCGGGTCTCGAGAGTTGACCAAGGTCAAGGACGTTTCAGATCTCCAGATCAC 685
Db 28 IleArgLeuLeuGlnLysPheAspGlnArgSerLysIlePheGluAenGluAenTyr 47
QY 686 CGTATTGTGAGATCGAGGCAACCTCAACACTCTTGTCTTCCCAAGCAGCGTGATGCT 745
Db 48 ArgLeuLeuGluTyrLysSerLysProHisThrIlePheLeuProGlnHisThrAspAla 67

QY 746 GATAACATCTTGTATTATCCAGCAAGGCAAGCCACCGTACCGCTAGCAATGCAATAAC 805
Db 68 AspTyrIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeuLysProAspAsp 87
QY 806 AGAAAGAGCTTTAAATCTTCAGCAGGCGCATCCACTCAGAAATCCATCCGCTTTCATTTC 865
Db 88 ArgAsnSerPheAsnLeuGluArgGlyAspThrIleLysLeuProAlaGlyThrIleAla 107
QY 866 TACATCTTGAACCGCCATGACCAACCAACCTCAGAGTAGCTAAAATCTCCATGCCGTT 925
Db 108 TyrLeuValAsnArgAspAsnGluGluLeuArgValLeuAspLeuAlaIleProVal 127
QY 926 AACACACCGCCGCTTTCAGGATTTCTTCCCGCGCAGCAGCGACCAATCATCTCTAC 985
Db 128 AsnArgProGlyGlnLeuGlnSerPheLeuLeuSerGlyAsnGlnAsnGlnAsnTyr 147
QY 986 TTGCAGGGCTTCAGCAGGAATACGTTGGAGGCGCTTCAATCGCGAAATTCATAGATA 1045
Db 148 LeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAspTyrGluGluIle 167
QY 1046 CGAGGGTCTCTTAGAAGAAATCGAGAGGTGAGCAAGAGAGAGAGGCGCAGAGCGGA 1105
Db 168 GluLysValLeuLeuGluHisGluLysGluThrGlnHisArgArgSerLeuLys--- 186
QY 1106 TGGAGTACTCGGAGTAGTACAGAACATGAGGAGTGTATGATCAAGTGTCAAGAGGAC 1165
Db 187 ---AspLysArgGlnGlnSerGlnGluGluAenValIleValLysLeuSerArgGlyGln 205
QY 1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGCTCAAGAAAGGCTCCGAAGAGAGGGA 1225
Db 206 IleGluLeuLeuSerLysAsnAlaLysSerThrSerLysLysSerValSerSerGluSer 225
QY 1226 GATATCACCAACCCCAATCAATTTGAGAGAGCGGAGCCGCTTTTCTTAACAATTTGG 1285
Db 226 Glu-----ProPheAsnLeuArgSerArgGlyProIleTyrSerAsnGluPheGly 242
QY 1286 AAGTTATTGAGGTGAAGCCAGCAAGAGAACACCCACCTTCAGACCTTGACATGATG 1345
Db 243 LysPhePheGluIleThrProGlu---LysAsnProGlnLeuGlnAspLeuAspIlePhe 261
QY 1346 CTCACCTGTGTAGATCAAGAGAGCTTTGATGCTCCACACCTTCAACTCAAGAGGCC 1405
Db 262 ValAsnSerValGluIleLysGluLysGluSerLeuLeuLeuProHisTyrAsnSerArgAla 281
QY 1406 ATGGTTATCGTCTGCTCAACAAAGAACTGGAAACCTTGAACCTCGTGGCTGTGAAGAA 1465
Db 282 IleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuValGlyGlnArgAsn 301
QY 1466 GAGCAACACAGAGGGGCGCGGGAAGAGAGGAGGAGCAAGACGAAAGAGAGGAGGA 1525
Db 302 GluAenGlnGlnGluGlnArgLysGluAspGluGluGluGlnGlnGlyGluGlu 321
QY 1526 AGTAACAGAGAGGTGCTAGGTACACAGCGAGGTGTAAGGAAGGCGCATGTTCATCATG 1585
Db 322 IleAsnLysGlnValGlnAsnTyrLysAlaLysLeuSerSerGlyAspValPheValIle 341
QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTGCTTCGGT 1645
Db 342 ProAlaGlyHisProValAlaLeuLysAlaSerSerAsnLeuAspLeuGlyPheGly 361
QY 1646 ATCAACGCTGAAAACCAACACAGATCTCTTCCTGAGGTGATAGGCAATGTGATAGAC 1705
Db 362 IleAsnAlaGluAenAsnGlnArgAsnPheLeuAlaGlyAspGluAspAsnValIleSer 381
QY 1706 CAGATAGAGCAAGCAAGGAGGATTTAGCATTTCCCTGGTCCGGTGAACCAAGTTGAGAAG 1765
Db 382 GlnValGlnArgProValLysGluLeuAlaPheProGlySerAlaGlnGlnValAspArg 401
QY 1766 CTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCAA 1825
Db 402 IleLeuGluAenGlnLysGlnSerHisPheAlaAspAlaGlnProGlnGlnArg----- 419

QY 1826 TCTCCGTCGTCCTCTGAGAAAGAGTCTCCTGAGAAAGAGGATCAA 1870
 Db 420 -----GluArgGlySerArgGluThrArgAspArg 429
 RESULT 7
 FWSYCB
 beta-conglycinin beta chain - soybean
 C:Species: Glycine max (soybean)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
 C:Accession: JQ0969
 R:Harada, J.J.; Barker, S.J.; Goldberg, R.B.
 Plant Cell 1, 415-425, 1989
 A:Title: Soybean beta-conglycinin genes are clustered in several DNA regions and are reg
 A:Reference number: JQ0969; PMID:93005638; PMID:2562562
 A:Accession: JQ0969
 A:Molecule type: DNA
 A:Residues: 1-439 <HAR>
 A:Cross-references: UNIPROT:P52974; GB:S44893; NID:g256426; PIDN:AAB23463.1; PID:g256427
 C:Comment: This protein accumulates during seed development and is hydrolyzed after germ
 C:Genetics:
 A:Gene: CG-4
 A:Introns: 101/1; 159/3; 186/3; 281/3; 375/1
 C:Superfamily: glycinin
 C:Keywords: seed; storage protein

Alignment Scores:
 Pred. No.: 4,95e-73 Length: 439
 Score: 1156.00 Matches: 237
 Percent Similarity: 69.98% Conservative: 87
 Best Local Similarity: 51.19% Mismatches: 85
 Query Match: 32.45% Indels: 54
 DB: 1 Gaps: 10

US-10-728-323-1 (1-2032) x FWSYCB (1-439)

QY 542 GTGAGGAGAAACATCTCGGAACAACCCCTTTCTACTTC---CCGTCAGGCGGTTTACG 598
 Db 26 ValArgGluAsp-----GluAsnAsnProPheTyrPheArgSerSerAsnSerPheGln 43
 QY 599 ACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTGCACCAAGGTCA 658
 Db 44 ThrLeuPheGluAsnGlnAsnValArgIleArgLeuGlnArgPheAsnIysArgSer 63
 QY 659 AGGCAGTTTCAGAACTCTCCAGAACTCACCGTATTGTGCAGATCGAGGCCAAACCTAACCT 718
 Db 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerIysProAsnThr 83
 QY 719 CTTGTGTTCTCCCAAGCAGCTGATGTGATTAACATCCTTGTATTCCAGCAAGGGCAAGCC 778
 Db 84 IleLeuLeuProHisIleAlaAspAlaAspPheLeuPheValLeuSerGlyArgAla 103
 QY 779 ACCGTGACCGTAGCAATGGCAATACAGAAAGAGCTTTAAATCTTCAGGCGCCATGCA 838
 Db 104 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 123
 QY 839 CTCAGAACTCCATCCGGTTTCATTCTCATCTTCAACCGCCATCACCAACCAAGACCTC 898
 Db 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
 QY 899 AGATGAGCTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCG 958
 Db 144 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 163
 QY 959 GCGAGCAGCGAGACCAATCATCTACTTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
 Db 164 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 183
 QY 1019 GCCTTCAATCGCGAAATTCAATGATGATACGAGGGTGCTGTGTAGAAGAGAATGCAGGAGT 1078
 Db 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 198
 QY 1079 GAGCAAGAGAGAGCGGCGAGCGGATGGAGTACTCGGAGTAGTGAGAACAAATGAGCA 1138
 Db 1079 GAGCAAGAGAGAGCGGCGAGCGGATGGAGTACTCGGAGTAGTGAGAACAAATGAGCA 1138

Qy	554	ACATCTCGAAACACCCCTTCTACTTCCGCTCAAGCGGTTTAGCACCCGCTACGGGAAC	6113
Db	30	SerAspGlnAspAsnProPheValPheGluSerAenArgPheGlnThrLeuPheGluAsn	49
Qy	614	CAAAACGTTAGATCCGGTCTCTCAGAGGTTTCACAAAGGTCAGCGAGTTTTCAGAAT	673
Db	50	GluAsnGlyHisIleArgLeuLeuGlnIysPheAspGlnHisSerIysLeuLeuGluAsn	69
Qy	674	CTCCAGAATACCCGPTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGTCTTCTCCCAAG	733
Db	70	LeuGlnAsnTyrArgLeuLeuGluTyrIysSerIysProHisThrIlePheLeuPheGln	89
Qy	734	CACGCTGATGCTGATAACACTCTTGTTATCCAGCAAGCGCAAGCCACCGTACCGTAGCA	793
Db	90	GlnThrAspAlaaspPheIleGluValValLeuSerGlyIysAlaIleLeuThrValLeu	109
Qy	794	AATGGCAATAACAGAAAGAGCTTTTAATCTTTGACGAGGGCCATGCATCAGAAATCCCATCC	853
Db	110	LeuProAsnAspArgAsnSerPheSerLeuGluArgGlyAspThrIleIysLeuProAla	129

QY	854	GGTTTCATTCTC	TACATCTTTGA	CCCGCCATGACA	ACCAGAACCTC	GAGTAGCTAA	AATC	913				
DB	130	GlyThrIleGly	TyrLeuValAsn	ArgAspGlu	AspLeuArg	ValLeuAsp	Leu	149				
QY	914	TCCATGCCGTT	TAACACACCC	CGGCGAGTT	TTCCTCCG	CGGAGCAGC	CGGAGAC	973				
DB	150	ValIleProVal	AsnArgProGly	GluProGln	SerPheLeu	LeuSerGly	AsnGlnAsn	169				
QY	974	CAATCATCTC	TACTTGCAGG	GCTTCAGCAG	GAATACTGTG	GAGGCGCTT	CAATCGCGAA	1033				
DB	170	GlnProSerIle	LeuSerGlyPhe	SerLysAsn	IleLeuGlu	AlaSerPhe	AsnThrAsp	189				
QY	1034	TTCATCAGAT	ACGAGGTCTG	TTTAGAGAG	AATCGCAG	GGTGACGAG	GAGAGAGA	1093				
DB	190	TyrlsGluIle	GluLysValLeu	LeuGluGlu	HisGlyLys	GluLysTyr	HisArgArg	209				
QY	1094	GGGCAGAGC	GATGAGTACT	CCGAGTAGT	GAGAACAT	CAAGCAG	GATGATGCA	AAAGTG	1153			
DB	210	GlyLeuLys	AspArgArg	GlnArgGly	GlnGluAsn	-----Val	IleValLysIle	227				
QY	1154	TCAAGAGC	ACGTTGAAG	ACTTACTTA	GACCGCTAA	ATCCGTCTC	CAAAAGAAG	CGCTCC	1213			
DB	228	SerArgLys	GlnIleGlu	LeuAsnLys	AsnAlaLys	SerSerLys	LysSerThr	247				
QY	1214	GAAGAAGC	GATATACCA	CCCCCAAT	CAACTTGAG	AGACGGC	CGCATCTT	CTTCT	1273			
DB	248	SerSerGlu	SerGlu-----	-----Pro	PheAsnLeu	LysSerArg	GluProIleTyr	Ser	264			
QY	1274	AACAAC	TTTCGGG	AAGTTATT	TGAGTCA	AGCCAGACA	GAAGAAC	CCCCCAGCTT	CAGGAC	1333		
DB	265	AsnLysPhe	GlyLysPhe	GluIleThr	Pro---Lys	ArgAsnPro	GlnLeuGln	Asp	283			
QY	1334	CTGCACAT	GATGCTC	ACTCTGT	AGATCAAA	AGAAGAG	CGCTTTG	ATGCTCC	CACACTTC	1393		
DB	284	LeuAsnIle	PheValAsn	TyrValGlu	IleAsnGlu	GlySerLeu	LeuLeuLeu	ProHisTyr	303			
QY	1394	AACTCAAG	CCCATGGT	TATCGTC	GTCTCAAC	AAAGGA	ACTGGAA	ACCTTGA	ACTCGTG	1453		
DB	304	AsnSerArg	AlaIleVal	IleValThr	ValAsnGlu	GlyLysGly	AspPheGlu	LeuVal	323			
QY	1454	GCTGTAAGA	AAAAAGCA	CAACAG	AGGGGAC	GGGGGAACA	GAAGGAG	GACGAAG	ACGAA	1513		
DB	324	GlyGlnArg	AsnGluAsn	GlnGlnGly	LeuArgGlu	GluTyrAsp	GluGluLys	GluGln	343			
QY	1514	GAAGAGAG	GGAAGTAAC	AGAGAG	GTGCTAG	TGTACAC	CGCGAG	TTTGAAG	GAAGCGAT	1573		
DB	344	GlyGluGlu	IleLutIleArg	LysGlnVal	GlnAsnTyr	LysAlaLys	LeuSerPro	GlyAsp	363			
QY	1574	GTGTTCA	TATGCAC	GAGCTCAT	CCAGTAG	CCATCA	CCGTTCT	CTCCGAA	CTCCATCTG	1633		
DB	364	ValLeuVal	IleProAla	GlyTyrPro	ValAlaIle	LysAlaSer	AsnLeuAsn	Leu	383			
QY	1634	CTTGGCTT	CCGTATCA	ACGCTG	AAAAACA	CCACAG	CAATCTT	CTCTG	CAGGTGATA	AGGAC	1693	
DB	384	ValGlyPhe	GlyIleAsn	AlaGlu	AsnAsnGln	ArgTyrPhe	LeuAlaGly	GluGluAsp	403			
QY	1694	AATGTGAT	AGACAGAT	AGAAAG	CAAGCA	AGGATTTAG	CATTCCT	CGGTG	CGGTGAA	1753		
DB	404	AsnValIle	SerGlnIle	HisLysPro	ValLysGlu	LeuAlaPhe	ProGlySer	AlaGln	423			
QY	1754	CAAGTT	CAGAAG	CTCATCA	AAAAAC	CGAAGAT	CTC	CAC	TTGTG	AGTGC	CTCCTCA	1813
DB	424	GluValAsp	ThrLeuLeu	GluAsnGln	LysGlnSer	HisPheAla	AsnAlaGln	ProGln	443			
QY	1814	TCTCAAT	CTCAAT	CTCGTCT	CTCTG	AGAAAG	AGTCTC	CTC	GAGAAAG	AGGATCA	AGAG	1873
DB	444	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	446
QY	1874	GAGGAA	AAAC	CAGAG	GGAAG	GGTCC	ACTCTCT	TTC	ATTTTGA	AGCTTTT	1924	
DB	447	ArgGlySer	GlnGluIle	LysAspHis	LeuTyrSer	IleLeuGly	SerPhe	463				

```

RESULT 10
S06309
vicilin precursor (clone VvVic1) - tick bean
N:Alternate names: 7S seed storage protein
C:Species: Vicia faba var. minor (tick bean)
C:Date: 31-Mar-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S06309; S06456
R:Weschke, W.; Baeumlein, H.; Wobus, U.
Nucleic Acids Res. 15, 10065, 1987
A:Title: Nucleotide sequence of a field bean (Vicia faba L. var. minor) vicilin gene.
A:Reference number: S06309; MUID:88096511; PMID:3697075
A:Accession: S06309
A:Molecule type: DNA
A:Residues: 1-463 <WES>
A:Cross-references: UNIPROT:P08438; EMBL:Y00506; NID:g829146; PIDN:CAAG8559.1; PID:g22051
A>Note: The authors translated the codon TCT for residue 296 as Cys
R:Weschke, W.; Bausuener, R.; van Hai, N.; Czibal, A.; Baeumlein, H.; Wobus, U.
Biochem. Physiol. Pflanz. 183, 233-242, 1988
A:Title: The structure of a Vicia faba vicilin gene.
A:Reference number: S06456
A:Accession: S06456
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-463 <WES2>
A>Note: 257-Lys and 443-His were also found
C:Genetics:
A:Introns: 102/1; 160/3; 187/3; 295/3; 400/1
C:Superfamily: glycinin
C:Keywords: seed; storage protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-463/Product: vicilin #status predicted <MAT>

Alignment Scores:
Pred. No.: 1,06e-69 Length: 463
Score: 1108.50 Matches: 218
Percent Similarity: 68.93% Conservative: 97
Best Local Similarity: 47.70% Mismatches: 119
Query Match: 31.12% Indels: 23
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S06309 (1-463)
Qy 554 ACATCTCGGAACAACCCCTTTCTACTCCCGTCGAAGCGGTTTAGCACCCGCTACGGGAAC 613
Db 30 SerAspGlnAspAsnProPheValPheGluSerAsnArgPheGlnThrLeuPheGluAsn 49
Qy 614 CAABAACGTAGATCCGGTCTCTGAGAGTTTGACCAAGGTCGAAGGCAGTTTCAGAAT 673
Db 50 GluAsnGlyHisIleArgLeuLeuGlnHisPheAspGlnHisSerLysLeuLeuGluAsn 69
Qy 674 CTCAGAAATCACCGGTATTGTGCAGATCGAGGCCAACCTTAACACTCTTGTTCTTCCCAAG 733
Db 70 LeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIlePheLeuProGln 89
Qy 734 CACGCTGATGCTGATAACATCTCTTATTCACAGCAAGGCCAAGCCACCGTCAGCGTAGCA 793
Db 90 GlnThrAspAlaAspPheIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeu 109
Qy 794 AATGCGAATAACAGAAAGAGCTTTTAATCTTGACGAGGCGCATGCATCAGAAATCCCATCC 853
Db 110 LeuProAsnAspArgAsnSerPheSerLeuGluArgGlyAspThrIleLysLeuProAla 129
Qy 854 GGTTCATCTTCATCATCTTGAAACCGCATGACAACCAACCTCAGATGAGTAAATC 913
Db 130 GlyThrIleGlyTyrLeuValAsnArgAspGluAspGluAspLeuArgValLeuAspLeu 149
Qy 914 TCCATGCCCGGTAAACACACCCGGCCAGTTTGAGGATTTCTCCCGCGCAGCACCCGAGAC 973
Db 150 ValIleProValAsnArgProGlyLupProGlnSerPheLeuLeuSerGlyAsnGlnAsn 169
Qy 974 CAATCATCTCTACTTCAGGGCTTCAGCAGGAATACGTTGGAGGCCCGCTTCAATGCGGAA 1033
Db 170 GlnProSerIleLeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAsp 189

```

```
QY 1034 TTCAATGAGATACGGAGGTGCTGTAGAGAGAAATGCAGGAGGTGAGCAAGAGGAGAGA 1093
Db 190 TyrlsGluileGluLysValLeuLeuGluGluHisGlyLysGluLysTyrlsHisArgArg 209
QY 1094 GGCAGAGCGGAGTACTCGAGTAGTGAGAACAAATGAAGGAGTGATAGTCAAAAGTG 1153
Db 210 GlyLeuLysAspArgGlnArgGlyGlnGluLysN-----ValleVallylle 227
QY 1154 TCAAGAGCAGCTTGAAGAACTTACTAGACGCTAAATCCCTCTCAAGAAAGGCTCC 1213
Db 228 SerArgLysGlnileGluLeuLeuAsnLysAsnAlaLysSerSerLysLysSerThr 247
QY 1214 GAAGAGCGGAGATATACCAACCCAACTCACTTGAGAGAGCGGAGCCGATCTTTCT 1273
Db 248 SerSerGluSerGlu-----ProPheAsnLeuArgSerArgGluProIleTyrlsSer 264
QY 1274 AACAACTTTGGGAGTTATTGTAGGTGAGCAGACAGAGAAAGACCCCGAGCTTCAAGGAC 1333
Db 265 AsnLysPheGlyLysPhePheGluileThrPro---LysArgAsnProGlnLeuGlnAsp 283
QY 1334 CTGGACATGATGCTACCTGTGTAGATCAAAAGAGGAGCTTTGATGCTCCACATCTC 1393
Db 284 LeuAsnilePheValAsnTyrlsValGluileAsnGluGlySerLeuLeuLeuProHisTyrls 303
QY 1394 AACTCAAGCCATGTTATCGTCTGCTCAACAAAGGAACTCGAAACCTTGAACCTCGTG 1453
Db 304 AsnSerArgAlaileValleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323
QY 1454 GCTGTGAAGAAAGAGCAACACAGAGGGGCGGAGAGAGAGAGAGAGAGAGAGAGAA 1513
Db 324 GlyGlnArgAsnGluAsnGlnGlnGlyLeuArgGluGluTyrlsAspGluGluLysGluGln 343
QY 1514 GAAGAGGAGGAGAGTAAACAGAGAGGTGCGTAGTACACAGCGAGGTGAAGGAGCGGAT 1573
Db 344 GlyGluGluGluileArgLysGlnValGlnAsnTyrlsAlaLysLysLeuSerProGlyAsp 363
QY 1574 GTGTTCAATGTCAGCAGCTCATCAGTAGCCATCAAGCTTCTCCGAACTCCATCTG 1633
Db 364 ValLeuValleProAlaGlyTyrlsProValAlaileLysAlaSerSerAsnLeuLeu 383
QY 1634 CTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTCTTCAGGTGATAGGAC 1693
Db 384 ValGlyPheGlylleAsnAlaGluAsnAsnGlnArgTyrlsPheLeuAlaGlyGluLysAsp 403
QY 1694 AATGTGATAGACCATAGAGAGCAAGCGAGGATTTAGCATTCCTCGGTGCGGTGAA 1753
Db 404 AsnValleSerGlnileHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423
QY 1754 CAAGTTGAGAGCTCATCAAAACCAAGAGGAACTCATCTTTGTGAGTGTCTGCTCTCAA 1813
Db 424 GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProArg 443
QY 1814 TCTCAATCTCAATCTCCGTCTCTCGTAGAAGAGTCTCTCGAGAAAGAGGATCAAGAG 1873
Db 444 -----GluArgGlu 446
QY 1874 GAGGAAACCAAGAGGAGGAGGTGCCACTCTTTCAATTTGAAGGCTTTT 1924
Db 447 ArgGlySerGlnGluileLysAspHisLeuTyrlsSerileLeuGlySerPhe 463
```

RESULT 11

FWPMVB
vicilin B precursor - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03344
R:Lycett, G.W.; Delauney, A.J.; Gatehouse, J.A.; Gilroy, J.; Croy, R.R.D.; Boulter, D.
Nucleic Acids Res. 11, 2367-2380, 1983
A:Title: The vicilin gene family of pea (Pisum sativum L.): a complete cDNA coding sequence
A:Reference number: A93462; MUID:83220791; PMID:6687941
A:Accession: A03344
A:Molecule type: mRNA

A;Residues: 1-410 <LYC>

A;Cross-references: UNIPROT:P02854

A;Experimental source: cv. Feltham First, clones pDUB7 and pDUB4

A>Note: parts of this sequence, including the amino end of the mature protein, were determined from the complementary DNA sequence of a cDNA library derived from a mature seedling of a multi-gene family coding for vicilin

C;Comment: The gene that codes for this protein is part of a multi-gene family coding for vicilin

C;Superfamily: glycoprotein

C;Keywords: glycoprotein

F;1-15/Domain: signal sequence #status predicted <SIG>

F;16-410/Product: vicilin type B (fragment) #status predicted <MAT>

F;321-322/Cleavage site: Asn-Asp (unidentified proteinase) #status experimental

F;359/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:

Pred. No.: 7,05e-67 Length: 410

Score: 1068.00 Matches: 209

Percent Similarity: 71.64% Conservative: 84

Best Local Similarity: 51.10% Mismatches: 104

Query Match: 29,98% Indels: 12

DB: 1 Gaps: 5

US-10-728-323-1 (1-2032) x FWPMB (1-410)

QY 542 GTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGCGGTTTAGCACC 601

Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32

QY 602 CGCTACGGGAACAAACCGGTAGGATCCGGGTCTCTGACAGAGTTTGACCAAGGTCAAGG 661

Db 33 LeuTyrlsGluAsnGluAsnGlyHisileArgLeuLeuGlnLysPheAspLysArgSerLys 52

QY 662 CAGTTTCAGAACTCCAGATCACCGTATTGTCAGATCGAGGCCCAACCTAACACTCTT 721

Db 53 IlePheGluAsnLeuGlnAsnTyrlsArgLeuLeuGluTyrlsSerLysProHisThrLeu 72

QY 722 GTTCTTCCCAAGCAGCTGCTGCTAATACATCTTGTATCCAGCAAGCGGCAAGCCACC 781

Db 73 PheLeuProGlnTyrlsThrAspAlaAspPheIleValLeuValLeuSerGlyLysAlaThr 92

QY 782 GTGACCGTAGCAAAATGGCAATACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCATC 841

Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaile 112

QY 842 AGAATCCCATCCGTTTCAATTTCTTCTACATCTTGACCGCATACACAGCAACCTCAGA 901

Db 113 LysLeuProAlaGlySerIleAlaTyrlsPheAlaAsnArgAspAsnGluGluProArg 132

QY 902 GTAGCTAAATCTCCATGCCGTAAACACACCCCGCCAGTTTGAGGATTTCTCCCGCG 961

Db 133 ValLeuAspLeuAlaileProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152

QY 962 AGCAGCGAGACCAATCATCTTCTTCAAGGCTTTCAGAGGAGTTCAGAGGAGGCGCC 1021

Db 153 GlyThrGlnAsnGlnLysSerSerLeuSerGlyPheSerLysAsnileLeuGluAlaAla 172

QY 1022 TTCATGCGGAATTCATGAGATACGGAGGTGCTGTAGAGAGAAATCCAGGAGGTGAG 1081

Db 173 PheAsnThrAsnTyrlsGluGluileGluLysValLeuLeuGluGlnGln-----Glu 189

QY 1082 CAAGAGGAGAGGCGCAGAGCGATCGGAGTACTCCGAGTAGTGAG---AACATGAAGGA 1138

Db 190 GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGluileAsnGluGluAsn 209

QY 1139 GTGATAGTCAAGTGTCAAGGAGCAGCGTTGAAGAACTTACTTAAGCAGCGCTAAATCCGTC 1198

Db 210 ValileVallyleValSerArgAspGlnileGluGluLeuSerLysAsnAlaLysSerSer 229

QY 1199 TCAAGAAAGGCTCCGAGAGAGAGGAGATATCAACCAACCCCAATCACTTGAGAGAGGC 1258

Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246

QY 1259 GAGCCCGCATCTTCTTAAACAACTTTGCGAAGTTATTGAGGTGAGCGAAGCAAGAGAAC 1318

Db 247 AsnProIleTyrlsSerAsnLysPheGlyLysPhePheGluileThrProGlu---LysAsn 265


```
QY 1319 CCCAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTGG 1378
Db 266 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysValGlySerLeu 285
QY 1379 ATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTGTCTCAACAAAGAACTGGA 1438
Db 286 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 305
QY 1439 AACCTTGAACCTCGTGGCTGTAGAAAGAGACAAACAGAGGGGCGGCGGAGAGAG 1498
Db 306 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 321
QY 1499 GAGGACGAAGCAAGAGAGGAGGAGAACTAACACAGAGAGTGGTAGGTACACAGCGAGG 1558
Db 322 AspLysGluGluGluGlnGluGluThrSerLysGlnValGlnLeuTyrArgAlaLys 341
QY 1559 TTGAAGGAGGCGATGTTTCATCATGCGCAGAGCTCATCCAGTACGCCATCAACGCTTCC 1618
Db 342 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 361
QY 1619 TCCGAACCTCATCTGCTTCCGCTATCATGCTGAAAGCAACACACAGATCTTCCTT 1678
Db 362 SerAspLeuAsnLeuIleGlyLeuGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeu 381
QY 1679 GCAGGTGATAAGGACAATGTGTATAGACCATAGAGAGCAAGCGAAGGATTTAGCATTC 1738
Db 382 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 401
QY 1739 CTGGGTCGGGTGAACAAGTTGAGAAG 1765
Db 402 ProGlySerSerHisGluValAspArg 410
RESULT 12
JQ2264
canavalin - jack bean
C:Species: Canavalia ensiformis (jack bean)
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: JQ2264; S19137
R:Ng, J.D.; Ko, T.P.; McPherson, A.
Plant Physiol. 101, 713-728, 1993
A:Title: Cloning, expression, and crystallization of jack bean (Canavalia ensiformis) ca
A:Reference number: JQ2264; MUID:94143475; PMID:8310055
A:Accession: JQ2264
A:Molecule type: mRNA
A:Residue: 1-445 <NGJ1>
A:Cross-references: UNIPROT:P50477; GB:X59467; GB:S76871; NID:g17976; PIDN:CAA42075.1; F
A:Experimental source: seed
R:Ng, J.D.; Stinchcombe, T.; Ko, T.P.; Alexander, E.; McPherson, A.
Plant Mol. Biol. 18, 147-149, 1992
A:Title: PCR cloning of the full-length cDNA for the seed protein canavalin from the jac
A:Reference number: S19137; MUID:92119225; PMID:1731967
A:Accession: S19137
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-445 <NGJ2>
A:Cross-references: EMBL:X59467; NID:g17976; PIDN:CAA42075.1; PID:g17977
C:Superfamily: glycinin
C:Keywords: storage protein
Alignment Scores:
Pred. No.: 1.74e-59 Length: 445
Score: 962.50 Matches: 198
Percent Similarity: 54.03% Conservative: 117
Best Local Similarity: 33.96% Mismatches: 101
Query Match: 27.02% Indels: 167
DB: 2 Gaps: 8
US-10-728-323-1 (1-2032) x JQ2264 (1-445)
QY 68 CCACGTGATCTTGTCTAGGATCTTGTCTGGCTTCAGTTCTTCGCAACGATGCCAAG 127
Db 8 ProLeuTrpLeuLeuGlyValValLeuLeuAlaSerValSerValSerValSerPheAlaHis 27
```

```
QY 128 TCATCACTTACCAGAGAAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
Db 28 Ser----- 28
QY 188 CAACAGGAAACCGATGACTTGAAGCAAAAGGATGCGAGTCTCGTGTGCACCAAGCTCGAG 247
Db 28 ----- 28
QY 248 TATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCCACCACCAACGTTCC 307
Db 29 -----GlyHisSerGly----- 32
QY 308 CCTCAGGCGGACGACACACGTGGCCGCAACCCGAGAGACTACGATGATGACCGCGTCAA 367
Db 33 -----GlyGluAlaGluAspGluSerGluGlu 41
QY 368 CCCGAAAGAGAGAGAGAGCGGATGGGGACAGCTGGACCGAGGGAGCGTGAAGAGAA 427
Db 42 SerArg----- 43
QY 428 GAAGACTGGAGACAACCAAGAGAGATTGGAGGCGCAAGTCTATCAGCAGCCCGGAAA 487
Db 43 ----- 43
QY 488 ATAAGGCCGAAGAGAGAGAGAACAGAGTGGGGAACACAGGTAGCATGTGAGG 547
Db 43 ----- 43
QY 548 GAAGAAACATCTCGAAACAACCTTTCTACTTCGCGTCAAGCGGCTTTAGCACCGCTAC 607
Db 44 -----AlaGlnAsnAsnProTyrLeupheArgSerAsnLysPheLeuThrLeuphe 60
QY 608 GGGACCAAAACGTTAGGATCGGGTCTCGAGAGTTTGACCAAAAGTCAAGGCAGTTT 667
Db 61 LysAsnGlnHisGlySerLeuLeuGlnArgPheAsnGluAspThrGlnLysLeu 80
QY 668 CAGAACTCTCAGAACTCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTCTTCT 727
Db 81 GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu 100
QY 728 CCCAAGACCGTGTATGCTGATTAACATCTTGTATTCAGCAAGGCGCAACCGTAC 787
Db 101 ProHisHisSerAspSerAspLeuValLeuValLeuGluGlyGlnAlaIleVal 120
QY 788 GTAGCAATGCGCAATAACAGAAAGAGCTTAACTTTGACAGAGGCGCATGCACTCAGAATC 847
Db 121 LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle 140
QY 848 CCATCCGCTTTTCATTTCTTACATCTTGAACCGCCATGACCAACCAAGACCTCAGAGTAGCT 907
Db 141 GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu 160
QY 908 AAAATCTCCATCGCCGTTAAACACCCGCGCAGTTTGTAGGATTTTCCCGCGAGAGCAGC 967
Db 161 LysPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerThr 180
QY 968 CGAGACCAATCATCTACTTCGCGGCTTCAGCAGGATACGTTGGAGGCGCGCTTCAAT 1027
Db 181 LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp 200
QY 1028 GCGGAATTCATATGATACGAGGAGTGTCTTTAGAGAGAATGACGAGGAGGTGAGCAAGAG 1087
Db 201 SerProTyrAspGluIleGluGlnThrLeuLeuGln----- 212
QY 1088 GAGAGAGGCGAGCGGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGAGTAGTAGTC 1147
Db 213 -----GluGluGlnGluGlyValIleVal 220
QY 1148 AAAGTGTCAAAGGACGAGTTCAGAACTTACTAAGCAGCGCTAAATCCGTCTCAAGAAA 1207
Db 221 LysMetProLysAspGlnIleGlnGluIleSerLysHisAlaGlnSerSerArgLys 240
```



```
QY 1208 GGCTCCGAAGAGGAGGATATACCAACCACTTGAAGAGCGGAGCCGAT 1267
Db 241 ThrLeuSerSerGln-----AspLysProPheAsnLeuArgSerArgAspProIle 257
QY 1268 CTTTCTAACAACTTTGGAGTTATTGAGTGAAGCCAGACAAAGAAACCCCGAGCTT 1327
Db 258 TyrSerAsnAsnTyrGlyLeuTyrGluLeuThrProGlu---LysAsnSerGlnLeu 276
QY 1328 CAGGACTGACATGATGCTCAGCTGTGTAGAGATCAAGAGAGCTTTGATGCTCCCA 1387
Db 277 ArgAspLeuAspIleLeuLeuAsnCysLeuGlnMetAsnGluGlyAlaLeuPheValPro 296
QY 1388 CACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGAACCTTGAA 1447
Db 297 HisTyrAsnSerArgAlaThrValIleLeuValAlaAsnGluGlyArgAlaGluValGlu 316
QY 1448 CTGCTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGAAAGAGGAGGACGAA 1507
Db 317 LeuValGlyLeuGluGlnGlnGln----- 326
QY 1508 GACGAAGAGAGGAGGGA---AGTAACAGAGAGGTGCTAGGTACACAGCGAGTTGAAG 1564
Db 327 -----GlyLeuGluSerMetGlnLeuArgArgTyrAlaAlaThrLeuSer 341
QY 1565 GAAGCGATGTGTTTCATCATCGCCAGCAGCTCATCCAGTACGCCATCAACGCTTCTCCGAA 1624
Db 342 GluGlyAspIleValIleProSerPheProValAlaLeuLysAlaAlaSerAsp 361
QY 1625 CTCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCGAGGT 1684
Db 362 LeuAsnMetValGlyIleGlyValAsnAlaGluAsnAsnGluArgAsnPheLeuAlaGly 381
QY 1685 GATAAGCAATGTGATAGACAGATAGAGAACAAAGCGAGGATTTAGCATTCCTCGGG 1744
Db 382 HisLysGluAsnValIleArgGlnIleProArgGlnValSerAspLeuThrPheProGly 401
QY 1745 TCGGGTGAACAAGTGTGAGAGCTCATCAAAACAGAGGAATCTCATTGTGAGTGCT 1804
Db 402 SerGlyGluGluValGluGluLeuLeuGluAsnGlnLysGluSerTyrPheValAspGly 421
QY 1805 CGTCTCCAA 1813
Db 422 GlnProArg 424

RESULT 13
S00281
C:Species: Canavalia gladiata (sword bean)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: S00281; S04598
R:Yamauchi, D.; Nakamura, K.; Asahi, T.; Minamikawa, T.
Eur. J. Biochem. 170, 515-520, 1988
A:Title: cDNAs for canavalin and concanavalin A from Canavalia gladiata seeds. Nucleotide
s.
A:Reference number: S00281; MUID:88111636; PMID:3338449
A:Accession: S00281
A:Molecule type: mRNA
A:Residues: 1-445 <YAM>
A:Cross-references: UNIPROT:P10562; EMBL:X06733; NID:g18003; PIDN:CAA29910.1; PID:g18004
R:Takei, Y.; Yamauchi, D.; Minamikawa, T.
Nucleic Acids Res. 17, 4381, 1989
A:Title: Nucleotide sequence of the canavalin gene from Canavalia gladiata seeds.
A:Reference number: S04598; MUID:89296493; PMID:2740227
A:Accession: S04598
A:Molecule type: DNA
A:Residues: 1-160, 'K', 162-445 <YAK>
A:Cross-references: EMBL:X15076; NID:g18006; PIDN:CAA333172.1; PID:g18007
C:Genetics:
A:Introns: 115/1; 173/3; 200/3; 290/3; 381/1
C:Superfamily: glycinin
C:Keywords: seed

Alignment Scores:
```

```
Pred. No.: 2,82e-59 Length: 445
Score: 59.50 Matches: 197
Percent Similarity: 54.03% Conservative: 118
Best Local Similarity: 33.79% Mismatches: 101
Query Match: 26.94% Indels: 167
DB: 2 Gaps: 8

US-10-728-323-1 (1-2032) x S00281 (1-445)
QY 68 CCATGATGCTGTGTAGGATCTCTGCTGGTCTAGTTTCTGCAACGATGCGCAAG 127
Db 8 ProLeuTyrLeuLeuLeuGlyValLeuLeuAlaSerValSerPheAlaHis 27
QY 128 TCATCACTTACCAGAGAAACAGAGAACCTCTGCGCCAGAGGTGCTCCAGAGTTGT 187
Db 28 Ser----- 28
QY 188 CAACAGAACCGATGACTTGAAGCAAAAGCATCGAGTCTCGTGTCCACCAAGCTCGAG 247
Db 28 ----- 28
QY 248 TATGATCTCTGTGTCTATGATCTCGAGGACACACTGGCACCAACCAACCAACGTTCC 307
Db 29 -----GlyHisSerGly----- 32
QY 308 CTTCCAGGGAGCGGACACAGCTGCGCCCAACCCGAGAGACTAGCATGATACCGCGCTCAA 367
Db 33 -----GlyGluAlaGluAspGluSerGluGlu 41
QY 368 CCCCAGAGAGAGAGGAGGCGGATGGGGACAGCTGGAGCCGAGGAGCGGTGAAGAGAA 427
Db 42 SerArg----- 43
QY 428 GAAGACTGGAGACAACCAAGAGAAGATTGGAGGCGCAAGTCTATCAGCAGCCACGGA 487
Db 43 ----- 43
QY 488 ATAAGCCCGAAGGAGAGAGAGAGAACAGAGTGGGGAAACACAGGTAGCCATGTGAGG 547
Db 43 ----- 43
QY 548 GAAGAAACATCTCGGAACAACCTTTCTACTTCCGTCAAGCGGTTTAGCACCCGCTAC 607
Db 44 -----AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe 60
QY 608 GGGAAACCAACCGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAGGCGAGTTT 667
Db 61 LysAsnGlnHisGlySerLeuArgLeuLeuGlnArgPheAsnGluAspThrGluLysLeu 80
QY 668 CAGATCTCCAGATCACCGTATTGTGAGATCGAGGCCAAACCTTACACTCTTGTCTTT 727
Db 81 GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu 100
QY 728 CCCAAGCAGCTGATGCTGATAAATCTTGTATTCTCAGCAAGGGCAAGCCCGTGACC 787
Db 101 ProHisHisSerAspSerAspLeuValLeuLeuGluGlyGlnAlaIleLeuVal 120
QY 788 GTAGCAATGGCAATACAGAAAGAGCTTTAATCTTTAGCAGGGCCCATCTCAGAAATC 847
Db 121 LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle 140
QY 848 CCATCCGGTTTCATTTCTTACATCTTTGAAACCCCATAGCAACACAGAACCTCAGAGT 907
Db 141 GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu 160
QY 908 AAAATCTCATCCCGTTAAACACACCGCCAGTTTGTAGGATTTCTTCCCGCGGAGCAGC 967
Db 161 AsnPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerSerThr 180
QY 968 CGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAATACGTTGGAGCCCGCTTCAAT 1027
Db 181 LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp 200
```

```
QY 1028 GCGGAATTCATAGATACGAGGGTGTCTTTAGAAAGATATGCAAGAGTGCAGCAAGAG 1087
Db   :::::::::::::::::::::
QY 201 SerProTyrAspGluIleGluGlnThrLeuLeuGln----- 212
Db   :::::::::::::::::::::
QY 1088 GAGAGAGGCGAGGCGGATCGAGTACTCGAGTAGTGAGAACATGAAGAGTGATAGTC 1147
Db   :::::::::::::::::::::
QY 213 -----GluGluGlnGluGlyValIleVal 220
Db   :::::::::::::::::::::
QY 1148 AAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTAAGCAGCGTAAATCCGCTCAAAGAA 1207
Db   :::::::::::::::::::::
QY 221 LysMetProLysAspGlnIleGlnGluIleSerLysHisAlaGlnSerSerArgLys 240
Db   :::::::::::::::::::::
QY 1208 GGCTCCGAAGAGAGGAGATATCAACCAACCCCAATCAACTTGAGAGAAAGCGCGCCGAT 1267
Db   :::::::::::::::::::::
QY 241 ThrLeuSerSerGln-----AspLysProPheAsnLeuArgSerArgAspProIle 257
Db   :::::::::::::::::::::
QY 1268 CTTTCTAAACAATTTCGGAGTTATTTAGGTGTGAAGCCAGACAGAAAGAACCCCGAGTT 1327
Db   :::::::::::::::::::::
QY 258 TyrSerAsnAsnTyrGlyLysLeuTyrGluIleThrProGlu---LysAsnSerGlnLeu 276
Db   :::::::::::::::::::::
QY 1328 CAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGAGGAGCTTTGATGCTCCA 1387
Db   :::::::::::::::::::::
QY 277 ArgAspLeuAspIleLeuLeuAsnCysLeuGlnMetAsnGluGlyAlaLeuPheValPro 296
Db   :::::::::::::::::::::
QY 1388 CACTTCAACTCAAAGGCCATGGTTATCGTCGTCTCAACAAAGGAAGAACTGGAAACCTTGA 1447
Db   :::::::::::::::::::::
QY 297 HisTyrAsnSerArgAlaThrValIleLeuValAlaAsnGluGlyArgAlaGluValGlu 316
Db   :::::::::::::::::::::
QY 1448 CTCGTGGCTGTAAAGAAAGAGCNAACAACAGAGGGGACGGCGGAGAGAGGAGACGAA 1507
Db   :::::::::::::::::::::
QY 317 LeuValGlyLeuGluGlnGlnGlnGln----- 326
Db   :::::::::::::::::::::
QY 1508 GACGAAGAAGAGGAGGGA---AGTACAGAGAGGTGCGTAGGTACACACGAGGTTGAAG 1564
Db   :::::::::::::::::::::
QY 327 -----GlyLeuGluSerMetGlnLeuArgTyrAlaAlaThrLeuSer 341
Db   :::::::::::::::::::::
QY 1565 GAAGCGGATGTGTTCATATGCGAGCAGCTATCCAGTAGCCATCAACGCTTCTCCCGAA 1624
Db   :::::::::::::::::::::
QY 342 GluGlyAspIleLeuValIleProSerSerPheProValAlaLeuLysAlaAlaSerAsp 361
Db   :::::::::::::::::::::
QY 1625 CTCATCTCTGCTGGCTTCGGTATCAAGCTGAAACACACAGAAATCTTCTCTCGAGGT 1684
Db   :::::::::::::::::::::
QY 362 LeuAsnMetValGlyIleGlyValAsnAlaGluAsnAsnGluArgAsnPheLeuAlaGly 381
Db   :::::::::::::::::::::
QY 1685 GATAAGCAATGTGTATAGACCATAGACAGAGCAAGCGAAGATTAGCATTCCTCTGG 1744
Db   :::::::::::::::::::::
QY 382 AsnLysGluAsnValIleArgGlnIleProArgGlnValSerAspLeuThrPheProGly 401
Db   :::::::::::::::::::::
QY 1745 TCGGTTGAACAGTTGAGAGCTCATCAAAAACCAAGAGAAATCTCAGTTTGTGAGTGTCT 1804
Db   :::::::::::::::::::::
QY 402 SerGlyGluGluValGluGluLeuLeuGluAsnGlnLysGluSerTyrPheValAspGly 421
Db   :::::::::::::::::::::
QY 1805 CGTCTCTCAA 1813
Db   :::::::::::::::::::::
QY 422 GlnProArg 424
Db   :::::::::::::::::::::
RESULT 14
T44430
protein PV100 [imported] - winter squash
C:Species: Cucurbita maxima (winter squash)
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44430
R:Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
J. Biol. Chem. 274, 2563-2570, 1999
A:Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single
A:Reference number: 222767; MUID:199107919; PMID:9891029
A:Accession: T44430
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-810 <YAM>
A:Cross-references: UNIPROT:Q92WI3; EMBL:AB019195; NID:93808061; PIDN:BAA34056.1; PID:93
Alignment Scores:
```

```
Pred. No.: 1.4e-55 Length: 810
Score: 907.50 Matches: 214
Percent Similarity: 54.65% Conservative: 121
Best Local Similarity: 34.91% Mismatches: 191
Query Match: 25.48% Indels: 87
DB: 2 Gaps: 20

US-10-728-323-1 (1-2032) x T44430 (1-810)

QY 188 CAACAGGACCGGAT---GACTTGAAGCAAAAGGATCGGAGTCTCGTGCACCAAGCTC 244
Db   :::::::::::::::::::::
QY 244 GlnArgAspProAspTyrArgArgGluGlnLysArgGluGlnGluGlnGluArg 263
Db   :::::::::::::::::::::
QY 245 GAGTATGATCCTCGTGTGTCTATCATCTCTGAGGACACACTGGCACCACCAACCAACGT 304
Db   :::::::::::::::::::::
QY 264 GluArgGluArgArgGlyGlyArgAspAspGluAspGlu-----AsnGlnArg 279
Db   :::::::::::::::::::::
QY 305 TCCCTCCAGGGAGCGGACACAGTCGCGGCCCAACCCCGAGAC---TACGATGATACCGC 361
Db   :::::::::::::::::::::
QY 280 AspProAspTyrArgGluGlnGluArgArgGluGlnGluArgArgArgArgGluArg 299
Db   :::::::::::::::::::::
QY 362 CGTCAACCCCGNAGAGAGAGGAGGCGGATGGGACCGAGCTGGACCGGAGGCGTGAA 421
Db   :::::::::::::::::::::
QY 300 GluGlnGluArgArgGluArgGluHisArgGlyGlyArgAspValGluAspGluAsnGln 319
Db   :::::::::::::::::::::
QY 422 AGAGAAGAAGACTGAGAGCAACCAAGAGAAGATTGGAGCGCAACCAAGTCATCAGCAGCCA 481
Db   :::::::::::::::::::::
QY 320 ArgAspProAspTyrArgGluGlnGluArgArgGluGlnGluArgGluGluGlnArgGlu 339
Db   :::::::::::::::::::::
QY 482 CGGAAA---ATAAGCCCGAGAGAGAGAGAAAGAGAAAGAGTGGGGAAACACCAAGTACG 538
Db   :::::::::::::::::::::
QY 340 ArgGluTyrGluArgGluHisGlyArgArgGlyArgGluGlnArgSerArgGluAsp 359
Db   :::::::::::::::::::::
QY 539 CATGTGAGG----- 547
Db   :::::::::::::::::::::
QY 360 GluArgArgArgHisGluArgGlnHisGlyArgSerArgValAsnGlnValAlaIle 379
Db   :::::::::::::::::::::
QY 548 -----GAAGAAACATCTCGGAACAACCTTCTACTTCCCTCAAGGCGGTTTACG 598
Db   :::::::::::::::::::::
QY 380 ArgArgThrGluGlnGlnSerAsnAsnProTyrTyrPheGlnGlnArgPheGln 399
Db   :::::::::::::::::::::
QY 599 ACCCGCTACGGAAACAAAACGGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCA 658
Db   :::::::::::::::::::::
QY 400 SerArgTyrArgSerAspGluGlyHisTyrArgValLeuGluArgPheSerGluArgSer 419
Db   :::::::::::::::::::::
QY 659 AGGCAAGTTTCAGATCTCCAGAAATCACCGTATTGTGCAGATCGAGCGCAACCACTAACACT 718
Db   :::::::::::::::::::::
QY 420 GluLeuLeuLysGlyIleLysAsnGlnArgLeuAlaLeuLeuGluAlaArgProHisThr 439
Db   :::::::::::::::::::::
QY 719 CTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCCAGCAAGGCAAGCC 778
Db   :::::::::::::::::::::
QY 440 PheIleValProHisHisLeuAspAlaGluCysValLeuLeuValValArgGlyArgAla 459
Db   :::::::::::::::::::::
QY 779 ACCGTGACCGTAGCA-----AATGGCAATAACAGAAAGAGCTTTAATCTTGAC 826
Db   :::::::::::::::::::::
QY 460 ThrIleThrValValGlnGlnLysArgGluThrArgLysGluSerTyrAsnValGlu 479
Db   :::::::::::::::::::::
QY 827 GAGGCGCATGCACTCAGAAATCCCATCCGTTTCTATTCCTACATCTTGAACCGCCATGAC 886
Db   :::::::::::::::::::::
QY 480 SerGlyAspValMetThrIleProAlaGlyThrThrLeuTyrLeuAlaAsnGln---Glu 498
Db   :::::::::::::::::::::
QY 887 AACCAAGAACCTCAGAGTAGCTAAATCTCCATGCGCCGTTAAACACACCGCGGCGAGTTTGAG 946
Db   :::::::::::::::::::::
QY 499 AsnGluAspLeuGlnIleValLysLeuValGlnProValAsnAsnProGlyGluPheLys 518
Db   :::::::::::::::::::::
QY 947 GATTTCTTCCCGGCGAGCGGAGACCAATCATCTTCTGAGGCGTTGACGAGGAAT 1006
Db   :::::::::::::::::::::
QY 519 AspTyrLeuSerAlaGlyGlyGluSerGln---AlaTyrTyrSerValPheSerAsnAsp 537
Db   :::::::::::::::::::::
QY 1007 ACGTTGAGGCGGCTTCAATCGGAATTCATGAGATACGAGGGGTCTGTTAGAGAG 1066
Db   :::::::::::::::::::::
QY 538 ValLeuGluAlaAlaLeuAsnIleProArgAspLysLeuGluArgIlePhe----- 554
Db   :::::::::::::::::::::
```

```
QY 1067 AATGCGAGGCTGAGCAAGGAGGAGGAGGCGGATGGAGTACTCGGAGTAGTGAG 1126
Db      ::::::::::: |||||
555 -----LysGlnArgArgGluArgGly----- 561

QY 1127 AACATGAAGAGGTAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTAAGCAC 1186
Db      ||| ::::::::::: |||||
562 -----GlyLysIleValArgAlaSerGlnGluGlnLeuArgAlaLeuSerGlnArg 578

QY 1187 GCTAATCCGTCTCAAGAAGAGGCTCGAAGAAGAGGAGATATCAACCCCAATCAAC 1246
Db      ||| ||||| ::::::::::: |||||
579 AlaThrSerVal---ArgLysGlySerArg-----GlyValArgAlaProIleLys 594

QY 1247 TTGAGAGAAGCGGAGCGGATCTTTCTAACAACCTTTGGGAAGTATTGAGGTGAAGCCA 1306
Db      ||| ||||| ::::::::::: |||||
595 LeuGluSerGlnThrProValTyrAsnAsnGlnTyrGlyGlnMetPheGluAlaCysPro 614

QY 1307 GACAAGAAGAACCCCGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAA 1366
Db      ||| ::| ||||| ::::::::::: |||||
615 Asp---GluPheProGlnLeuArgArgThrAspValAlaThrSerValValAspIleLys 633

QY 1367 GAAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGCTGCTCAAC 1426
Db      ::||| ::| ||||| ::::::::::: |||||
634 GlnGlyGlyMetMetValProHisPheAsnSerArgAlaThrTrpValValPheValSer 653

QY 1427 AAAGGAACCTGGAACCTTGAACCTCGTGCT-----GTAAGAAAGAGCAACAACAGAG 1480
Db      ::||| ||||| ::::::::::: |||||
654 GluGlyAlaGlySerPheGluMetAlaCysProHisIleGlnSerSerGlnTrpGlnArg 673

QY 1481 GGACGGCGGAAGAAGAG-----GAGGACGAAGCAAGAAAGAGGAGGA 1525
Db      ||| ||||| ::::::::::: |||||
674 GlyArgGluGluGluGluArgHisTrpArgArgGluGluGluGluGluGluGluArg 693

QY 1526 AGTAACAGAGAGTGGGTAGGTACACAGCGAGTTGAAGAGCGCATGTTTCATCATG 1585
Db      ||| ||||| ::::::::::: |||||
694 SerGlyArg---PheGluArgValAlaGlyArgLeuSerGluGlyGlyValLeuValIle 712

QY 1586 CCAGCAGCTCATCCGCTAGCATCAACGCTTCC-----TCCGAACCTCCATCTGCTGGC 1639
Db      ::||| ||||| ::::::::::: |||||
713 ProAlaGlyHisProIleAlaIleMetAlaSerProAsnGluAsnLeuArgLeuValGly 732

QY 1640 TTCGGTATCAACGCTGAAACCAACCAAGATCTCTTTCGAGGTGATTAAGACATGTG 1699
Db      ::||| ||||| ::::::::::: |||||
733 PheGlyIleAsnAlaGluAsnAsnHisArgAsnPheLeuAlaGly---ArgGluAsnIle 751

QY 1700 ATAGACCAGATAGAAAGCAAGCAAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAGTT 1759
Db      ::||| ||||| ::::::::::: |||||
752 MetAsnGluLeuAspArgGluAlaLysGluLeuAlaPheAsnValGluGlyLysGlnAla 771

QY 1760 GAGAACTCATCAAAACCAAGGAATCTCACTTTGTGTGCTGCTCTCAATCTCAA 1819
Db      ::||| ||||| ::::::::::: |||||
772 AspGluIlePheArgSerGlnArgGluSerPhePhe----- 783

QY 1820 TCTCAATCTCCGCTGCTCTCTGAGAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAA 1879
Db      ::||| ||||| ::::::::::: |||||
784 ThrGluGlyProGluGlyGlyArgArgSerThrGlu----- 796

QY 1880 AACCAAGGAGGAGGCTCACTCTCTTCAATTTTGAAG 1918
Db      ::| ::| ||||| ::::::::::: |||||
797 -----ArgSerProLeuLeuSerIleLeuLys 805
```

RESULT 15

S06398
alpha-globulin type A precursor - upland cotton
N;Alternate names: seed storage protein
C;Species: Gossypium hirsutum (upland cotton)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: S06398
R;Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
A;Reference number: S06398
A;Accession: S06398
A;Status: not compared with conceptual translation

A:Molecule type: DNA
A;Residues: 1-605 <CHL>
A;Cross-references: UNIPROT:P09799
C;Superfamily: glycinin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-605/Product: alpha-globulin type A #status predicted <MAT>

Alignment Scores:

Pred. No.: 8,2e-51 Length: 605
Score: 839.00 Matches: 216
Percent Similarity: 50.86% Conservative: 108
Best Local Similarity: 33.91% Mismatches: 223
Query Match: 23.55% Indels: 90
DB: 2 Gaps: 21

US-10-728-323-1 (1-2032) x S06398 (1-605)

```
QY 65 TCTCCACTGATGCTGTGTCTAGGATCTCTTCTGGCTTCAGTTTCTGCAAGCATGCC 124
Db      ||| ::| ||||| ::::::::::: |||||
6 SerValPheValValLeuLeuPheSerLeuPheGlyLeuLeuCysSerAla 25

QY 125 AAGTCATCATCTTACCAGAGAAACAGAGAACCCCTGCGCCAGAGG-----TGC 175
Db      ||| ::| ||||| ::::::::::: |||||
26 LysAspPheProGlyArgArgSerGluAspAspPro---GlnGlnArgTyrGluAspCys 44

QY 176 CTCGAGAGTTGTCAACAGGAACCG---GATGACTTGAAGCAAAAGCATGCGAGTCTCC 232
Db      ||| ::| ||||| ::::::::::: |||||
45 ArgLysArgCysGlnLeuGluThrArgGlyGlnThrGluGlnAspLysCysGluAspArg 64

QY 233 TGC---ACCAAGCTCGAG-----TATGATCCTCGTTGTGTCTAT 268
Db      ||| ::| ||||| ::::::::::: |||||
65 SerGluThrGlnLeuLysGluGlnGlnArgAspGlyGluAspProGlnArgArgTyr 84

QY 269 ---GATCTCGAGGACACATGCGCACCAACCAACAGTTTCCCTCCAGGG----- 316
Db      ||| ||||| ::::::::::: |||||
85 GlnAspCysArgGlnHisCysGlnGlnGluGluArgArgLeuArgProHisCysGluGln 104

QY 317 -----GAGCGGACACGTGCGCCCAACCCGAGAGTACGATGATGAC 358
Db      ||| ::| ||||| ::::::::::: |||||
105 SerCysArgGluGlnTyrGluLysGlnGlnGlnGlnPro-----Asp 119

QY 359 CCGCTGCAACCCGAGAGAGAGGAGCGCGATGCGGACCAAGCTGGACCG----- 409
Db      ::||| ||||| ::::::::::: |||||
120 LysGlnPheLysGluCysGlnGlnArgCysGlnTrpGlnGluGlnArgProGluArgLys 139

QY 410 -----AGGAGCGTGAAGAGAGAGAGACTGGAGACAAACCAAGA 448
Db      ||| ||||| ::::::::::: |||||
140 GlnGlnCysValLysGluCysArgGluGlnTyrGlnGluAspProTrpLysGlyGluArg 159

QY 449 GAAGAT---TGGAGCGGACCAAGTCATCAGCAGCCAGCAAAATAAGGCCCAAGGAAGA 505
Db      ||| ||||| ::::::::::: |||||
160 GluAsnLysTrpArgGluGluGluGlu-----GluSerAsp 172

QY 506 GAAGGAGAACAGAGTGGGGAACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAAC 565
Db      ||| ||||| ::::::::::: |||||
173 GluGlyGluGlnGln-----GlnArgAsn 180

QY 566 AACCCCTTTCTACTCCCGTCAAGCGGTTTATAGCACCCGCTACGGGAACCAAAACGCTAG 625
Db      ||| ||||| ::::::::::: |||||
181 AsnProTyrTyrPheHisArgArgSerPheGlnGluArgPheArgGluGluHisGlyAsn 200

QY 626 ATCCGGGTCTGCGAGAGTTTGACCAAAAGGTCAAGCGAGTTTCAGAAATCTCAGAAATC 685
Db      ||| ||||| ::::::::::: |||||
201 PheArgValLeuGlnArgPheAlaAspLysHisHisLeuLeuArgGlyIleAsnGluPhe 220

QY 686 CGTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTGTCTTCCCAAGCAGCCTGATGCT 745
Db      ||| ||||| ::::::::::: |||||
221 ArgIleAlaIleLeuGluAlaAsnProAsnThrPheValLeuProHisHisCysAspAla 240

QY 746 GATAACATCTTGTATTATCCAGCAAGGCAAGCCAGCCGTCAGCAAAATGCCAATAAC 805
Db      ||| ||||| ::::::::::: |||||
241 GluLysIleTyrValValThrAsnGlyArgGlyThrValThrPheValThrHisGluAsn 260
```

Search completed: August 24, 2005, 10:06:34
Job time : 54.4416 secs

```
QY 806 AGAAGAGCTTTAACTTCACGAGGGCCATGCACCTCAGAAATCCCATCCGTTTCATTTCC 865
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
261 LysGluSerTyrAsnValValProGluValValValArgIleProAlaGlySerThrVal 280
QY 866 TACATCTTGAACCGCCATGACAAACCACTCAGAGTAGCTAAATCTCCATGCCCGTT 925
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
281 TyrLeuAlaAsnGlnAspAsnArgGluLysLeuThrIleAlaValLeuHisArgProVal 300
QY 926 AACACACCCGGCCAGTTTCAGGATTTCTCCCGGCGAGCGAGACCAATCATCCTAC 985
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
301 AsnAsnProGlyGlnPheGlnLysPhePheProAlaGlyGlnGluAsnProGlnSerTyr 320
QY 986 TTGACGGCTTCAGCAGGAATACGTTGGAGCGCCCTTCAATCGCGAATTCATGAGATA 1045
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
321 LeuArgIlePheSerArgGluIleLeuGluAlaValPheAsnThr----- 335
QY 1046 CGGAGGGTCTGTTAGAAGAGATGCGAGAGGTGAGCAAGAGAGAGAGGGCAGAGCGGA 1105
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
336 ArgSerGluGlnLeuAspGluLeuProGlyIleArgGlnSerHisArgArgGlnGln--- 354
QY 1106 TGGAGTACTCGGAGTAGTCAGAACATGAAGAGAGTAGTCAAGTGTCAAAGGAGCAC 1165
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
355 -----GlyGlnGlyMetPheArgLysAlaSerGlnGluGln 366
QY 1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGGA 1225
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
367 IleArgAlaLeuSerGlnGlyAlaThrSerProArgGlyLysGlySerGluGlyTyrAla 386
QY 1226 GATATCACCACCAATCAACTTGAGAGAGGGGAGCCGATCTTTCTAACAACTTTGGG 1285
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
387 -----PheAsnLeuLeuSerGlnThrProArgTyrSerAsnGlnAsnGly 401
QY 1286 AAGTTATTGTGAGTGAAGCAGACAGAAAGAACCCCGCTTCAGGACCTGGACATGATG 1345
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
402 ArgPheTyrGluAlaCysProArgAsnPheGlnGlnGlnLeuArgGluValAspSerSer 421
QY 1346 CTCACCTGTGTAGAGATCAAAGAGAGGCTTTGATGCTCCACACTTCAACTCAAAGGCC 1405
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
422 ValValAlaPheGluIleAsnLysGlySerIlePheValProHisTyrAsnSerLysAla 441
QY 1406 ATGGTTATCGTCTGCTCAACAAAGAACTGAAACCTTGAACCTCGTGGCT-----GTA 1459
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
442 ThrPheValValLeuValThrGluGlyAsnGlyHisValGluMetValCysProHisLeu 461
QY 1460 AGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGAGGACGAGACCAAGAGAGAG 1519
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
462 SerArgGlnSerSerAspTrpSerSerArgGluGluGluGlnGluGluGluVal 481
QY 1520 GAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAAGCGGATGTTC 1579
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
482 GluArgArgSerGlyGlnTyrLysArgValArgAlaGlnLeuSerThrGlyAsnLeuPhe 501
QY 1580 ATCATGCCAGCAGCTCATCCAGTAGCCATCAAGCTTCC-----TCGAACTCCATCTG 1633
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
502 ValValProAlaGlyHisProValThrPheValAlaSerGlnAsnGluAspLeuGlyLeu 521
QY 1634 CTTGGCTTCGGTATC---AACGCTGAAACCAACACAGAAATCTTCTTCGACGTCGATAG 1690
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
522 LeuGlyPheGlyLeuThrAsnGlyGlnAspAsnLysArgIlePheValAlaGly---Lys 540
QY 1691 GACAATGTGATACACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGTCGGGT 1750
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
541 ThrAsnAsnValArgGlnTrpAspArgGlnAlaLysGluLeuAlaPheGlyValGluSer 560
QY 1751 GAACAAGTTGAGAGCTCATCAAAACACAG---AAGAACTCTCACTTTGTGAGTCTGCT 1807
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
561 ArgLeuValAspGluValPheAsnAsnAsnProGlnGluSerTyrPheValSerGlyArg 580
QY 1808 -----CCTCAATCTCAATCTCAATCTCGCTCGTCTCCT 1840
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
581 AspArgArgGlyPheAspGluArgGlySerAsnAsnProLeuSerPro 597
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 124.355 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3562
Sequence: 1 aataacatataattcatc.....cgttgtggtgtttcttcc 2032

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model
-Q=/cgn2 1/USPTO.spool/US10728323/runat_23082005_124355_29210/app_query.fasta.1.4757
-DB=Uniprot_03 -Qfmt=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10728323 @CGN 1.1.291 @runat_23082005_124355_29210 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	92.3	626	1 AL12 ARAHY	P43238 arachis hyp
2	3041	85.4	614	1 AL11 ARAHY	P43237 arachis hyp
3	2887.5	81.1	580	2 O6PSU3	O6psu3 arachis hyp
4	2172.5	61.0	428	2 O6PSU4	O6psu4 arachis hyp
5	1495.5	42.0	299	2 O6PSU5	O6psu5 arachis hyp
6	1435	40.3	303	2 O6PSU6	O6psu6 arachis hyp
7	1302	36.6	621	2 O7XXT2	O7xt2 glycine max
8	1290	36.2	533	2 O6EBC1	O6ebc1 lupinus alb
9	1284	36.0	621	2 O948Y0	O948y0 glycine max
10	1282	36.0	605	2 O94LX2	O94lx2 glycine max
11	1281	36.0	605	1 GLCA SOYBN	P13916 glycine max
12	1267	35.6	571	1 CVCA PEA	P13915 pisum sativ
13	1265.5	35.5	623	2 O948X9	O948x9 glycine max
14	1258.5	35.3	545	2 O41674	O41674 vicia narbo
15	1251	35.1	613	2 O9M3X6	O9m3x6 pisum sativ
16	1251	35.1	639	1 GLC4 SOYBN	P11827 glycine max

17	1190	33.4	559	2 O9FZP9	O9fzp9 glycine max
18	1183.5	33.2	418	2 O840U1	O84u1 lens culina
19	1180.5	33.1	543	2 O22120	O22120 glycine max
20	1178.5	33.1	459	1 VCLC PEA	P13918 pisum sativ
21	1168	32.8	415	2 O84U10	O84u10 lens culina
22	1165	32.7	416	2 O22121	O22121 glycine max
23	1156	32.5	439	1 GLCB SOYBN	P25974 glycine max
24	1152	32.3	439	2 O93VL9	O93vl9 glycine max
25	1148	32.2	438	2 O43626	O43626 pisum sativ
26	1145	32.1	415	2 Q702P1	Q702p1 pisum sativ
27	1135	31.9	415	2 Q702P0	Q702p0 pisum sativ
28	1123.5	31.5	463	2 O41677	O41677 vicia narbo
29	1108.5	31.1	463	1 VCL VICFA	P08438 vicia faba
30	1082	30.4	396	2 O84UB3	O84ub3 glycine max
31	1068	30.0	410	1 VCLB PEA	P02854 pisum sativ
32	1030	28.9	518	2 O9M3X8	O9m3x8 lens culina
33	962.5	27.0	445	1 CANA CANEN	P50477 canavalia e
34	959.5	26.9	445	1 CANA CANGL	P10562 canavalia g
35	910	25.5	593	2 O9SEW4	O9sew4 juglans reg
36	907.5	25.5	810	2 O9ZWI3	O9zwi3 cucurbita m
37	880	24.7	481	2 Q7YIC1	Q7yic1 juglans nig
38	865.5	24.3	666	2 Q9SPL4	Q9spl4 macadamia i
39	863.5	24.2	625	2 Q9SPL3	Q9spl3 macadamia i
40	849.5	23.8	666	2 Q9SPL5	Q9spl5 macadamia i
41	839	23.6	605	1 VCLA GOSHI	P09799 gossypium h
42	825.5	23.2	436	1 PHSA PHAVU	P07219 phaseolus v
43	806.5	22.6	430	2 Q41115	Q41115 phaseolus v
44	805.5	22.6	430	2 Q43633	Q43633 phaseolus v
45	804	22.6	423	1 PHS2 PHALU	Q43617 phaseolus l

ALIGNMENTS

RESULT 1

ID AL12 ARAHY STANDARD; PRT; 626 AA.
AC P43238;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Allergen Ara h 1, clone P418 precursor (Ara h 1).
OS Arachis hypogaea (peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Florunner;
RX MEDLINE=96013631; PubMed=7560062;
RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;
RT "Recombinant peanut allergen Ara h I expression and IGE binding in
patients with peanut hypersensitivity";
RL J. Clin. Invest. 96:1715-1721(1995).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE ASN-521.
RX MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;
RA Kolarich D., Altmann F.;
RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass
spectrometry of electrophoretically separated nonmammalian proteins:
application to peanut allergen Ara h I and olive pollen allergen Ole e
1.";
RL Anal. Biochem. 285:64-75(2000).
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)

or send an email to license@isb-sib.ch.

CC -----
 CC EMBL; L34402; AAB00861.1; -;
 DR HSSP; P25974; 11PJ;
 DR InterPro; IPR06045; Cupin.
 DR InterPro; IPR007113; Cupin_region.
 DR InterPro; IPR011051; RmlC_like_cupin.
 DR Pfam; PF00190; Cupin; 2.
 KW Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 626 Allergen Ara h 1, clone P41B.
 FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 626 AA; 71345 MW; 1A6BBB41490D0E3 CRC64;

Alignment Scores:

Pred. No.: 2 34e-207 Length: 626
 Score: 3286.00 Matches: 626
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.25% Indels: 0
 DB: 1 Gaps: 0

US-10-728-323-1 (1-2032) x AL12_ARAHY (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTGTCTCGCTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGATGCCAAGTCATCACCTTACCAGAGAAAAACAGAACCCCTGCGCCGAG 169
 DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTCAAGCAAAAGCATCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTGCTATGATCTCGAGGACACACTGGC 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACAGCTGGCGGCCCAACCCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGCTCAACCCGAAAGAGAGGAGGCGGATGGGGACAGCTGGACCG 409
 DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAAGAGAAGACTGGAGACCAACCAAGAGAAGATTGGAGGCGCAAGT 469
 DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
 QY 470 CATCAGCACCCCGAATAGGCCGGAAGGAAGAGAGAGAACAGAGTGGGGGAACA 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnIleArgGlyThr 160
 QY 530 CAGGTAGCATGTAGGAAGAAACATCTCGAAACACCCCTTCTACTTCCGCTCAAGG 589
 DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTACACCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCGAGAGGTTTGAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAGGCAGTTTCAGATCTCCAGATCACCGTATTGTGCAGATCCAGGCCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGTAGTAACATCTCTGTTATCCAGCAA 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCCGCTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTTAATCTTGACGAG 829

DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCATGCACTCAGAAATCCCATCCGTTTCATTTCTACATCTTGAACCGCCATGACAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTCAGGAT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTTCCCGCGAGCAGCGAGACCAATCTACTTTCAGGGCTTCAGCAGGAATAGC 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGAGGCGGCTTCAATGCCGAATTCATAGATACCGAGGCTCTGTAGAGAGAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluAsn 340
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGGAGCGGATGGAGTACTCGAGTAGTGAGAAC 1129
 DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGGAGTATAGTCAAAAGTGTCAAAGGAGCAGTTGAAGAACTTACTAAGCAGCT 1189
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCATCACTTG 1249
 DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAAGCGAGCCGATCTTTAAACAATTTGGGAAGTTATTGAGTGAAGCCAGAC 1309
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAAGACCCAGCTTCAGACCTGCATGATGCTCACCTGTGTAGAGATCAAGAA 1369
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440
 QY 1370 CGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAATCGAAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGAGCGGG 1489
 DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArg 480
 QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAAGAGGGAAGTAACAGAGAGGTGCTAGGTAC 1549
 DB 481 GluGluGluGluAspGluAspGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGTTCAAGAGGAGCGATGTTTCATGCCAGAGCTCATCCAGTACCCATC 1609
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACCTCATCTGCTTGGCTTCCGTTATCAACGCTGAAACCAACACAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
 QY 1670 ATCTTCTTGCAGGTGATAAGCAATGTGTAGACCAAGATAGAGAACAAGCGAGGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTGGGTGGGTGAAACAGTTGAGAGCTCATCAAAACACAGAGGAATCT 1789
 DB 561 LeuAlaPheProLysSerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCGTCTCTCTCTGAGAAAGAG 1849
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerProSerProSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAAAAACAGAGGAGGAGGTCTCTCTTCA 1909

```

Db      601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY      1910 ATTTTGAAGCGTTTAAAC 1927
Db      621 IleLeuLysAlaPheAsn 626
RESULT 2
ID11_ARAHY STANDARD; PRT; 614 AA.
AC P43237;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Allergen Ara h 1, clone P17 precursor (Ara h 1).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Florunner;
RX MEDLINE=96013631; PubMed=7560062;
RA Burke A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;
RT "Recombinant peanut allergen Ara h 1 expression and IgE binding in
RT patients with peanut hypersensitivity.";
RL J. Clin. Invest. 96:1715-1721 (1995).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE ASN-516.
RX MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;
RA Kolarich D., Altmann F.;
RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass
RT spectrometry of electrophoretically separated nonmammalian proteins:
RT application to peanut allergen Ara h 1 and olive pollen allergen Ole e
RT 1.";
RL Anal. Biochem. 285:64-75 (2000).
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L38853; AAA60336.1; --
DR HSSP; P25974; LIPJ.
DR GlycoSuiteDB; P43237; --
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR Pfam; PF00190; Cupin; 2.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 614 Allergen Ara h 1, clone P17.
FT CARBOHYD 516 516 N-linked (GlcNAc...)/FTId=CAR_000218.
FT SEQUENCE 614 AA; 70283 MW; 1DDACF217EBC5F31 CRC64;
Alignment Scores:
Pred. No.: 2,9e-191 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 85.37% Indels: 18
DB: 1 Gaps: 7
US-10-728-323-1 (1-2032) x AL11_ARAHY (1-614)
QY 50 ATGAGAGGGGGTTTCTCCATGATGCTGTGCTAGGATCCTTCTCGCTTCAGTT 109

```

```

Db      1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY      110 TCTGCAACGCATGCCAAGTCATCACCTTACAGAGAAAACAGAGAACCCCTGCGCCAG 169
Db      21 SerAlaThrGlnAlaLys--SerProTyr--ArgLysThrGluAsnProCysAlaGln 38
QY      170 AGTGCTCTCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTCT 229
Db      39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY      230 CGCTGCACCAAGCTCGAGTATGATCTCGTGTGTCTATGATCTCTCGAGGACACACTGGC 289
Db      59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY      290 ACCCAACCAACGTTCCCTCCAGGGAGCGGACACGTTGGCGGCCCAACCCGGAGACTAC 349
Db      75 AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY      350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGAGGAGCGGATGGGGACGAGCTGGACCG 409
Db      95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY      410 AGGAGCGGTGAAGAGAGAAAGACTGGAGACAACCAAGAGAGAAATTGGAGGCGACCAAGT 469
Db      115 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY      470 CATCAGACGCCACGGAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGGAACA 529
Db      135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGlnTyrGlyThr 154
QY      530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGAACAACCCCTTCTACTCCGCTCAAGG 589
Db      155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY      590 CGTTTAGCACCGCTACGGGAACCAAAACGCTAGGATCCGGTCTCTGAGAGTTTGAC 649
Db      175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgLeuValLeuGlnArgPheAsp 194
QY      650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAA 709
Db      195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY      710 CTTAACAATCTTGTCTTCCCAAGCAGCTGATGATGATAACAATCTTGTATTATCGAGAA 769
Db      215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY      770 GGGCAAGCCACCGTCACCGTAGCAATGCAATAACAGAAAGAGCTTTTAATCTTGACGAG 829
Db      235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY      830 GGCCATGCATCAGAAATCCCATCCCGTTTTCATTTCTTCTACATCTTGAACCGCCATGACAAC 889
Db      255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY      890 CAGAACCTTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCGCAGTTTGAAGAT 949
Db      275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY      950 TTCTTCCCGCGAGCAGCGGAGACCAATCATCTACTTTCAGGGCTTCAGAGGAGTACG 1009
Db      295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY      1010 TTGGAGGCGCGCTTCAATGCGGAATTCATATGAGATACGAGGGGTGCTGTTAGAAGAGAT 1069
Db      315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
QY      1070 GCAGAGGTGACCAAGAGAGAGAGGAGCGGAGCGGATGAGTACTCGGAGTAGTAGAAC 1129
Db      335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY      1130 AATGAGGAGTAGTAGTCAAGTGTCAAGGAGCAGTTGAAGAGACTTACTTAAGCAGCT 1189

```

Db 354 AsnGluGlyValIleValIysValSerIysGluHisValGlnLeuThrIysHisAla 373
QY 1190 AAATCGCTCTCAAGAAAGCGTCCGAAAGAGGAGATATCACCAACCAATCAACTTG 1249
Db 374 LysSerValSerIysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAGAAGCGACCGCATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 393 AsgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValIysProAsp 412
QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGGAGATGCTCACCTGTGTAGATCAAGAA 1369
Db 413 LysIysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleIysGlu 432
QY 1370 GGAGCTTTTCATGCTCCACACTCAACTCAAGGCCATGTTATCGTCTGCTCAACAAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerIysAlaMetValIleValValAsnIys 452
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGAGCGCG 1489
Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgIysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGAGACGACGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
Db 473 GluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGluGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTGAAGAGCGGATGTTTCATCATGCGCAGCAGCTCATCCA 1600
Db 493 ArgArgTyThrAlaArgLeuIysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGGCATCAACGCTTCTCGAATCTCATCTGCTTGGCTTCGGTATCAACGCTCAAAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCCAGAGATCTTCTTCAGGTGATAGGACCAATGTGATGACAGATAGAGCA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspIysAspAsnValIleAspGlnIleGluIysGln 552
QY 1721 GCGAAGGATTTAGCATTTCCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACCAAG 1780
Db 553 AlaIysAspLeuAlaPheProGlySerGlyGluGlnValGluIysLeuIleIysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCGCTCTCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAGAGTCTCCTGAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
Db 589 -----SerProGluIysGluAspGlnGluGluGluAsnGlnGlyGlyIysGlyPro 605
QY 1901 CTCCTTTCAATTTTGAAGGCTTTTAAAC 1927
Db 606 LeuLeuSerIleLeuIysAlaPheAsn 614

RESULT 3
Q6PSU3
ID AC Q6PSU3 PRELIMINARY; PRT; 580 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Li H., Wang L., Liao B., Yan Y., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581852; AAT00597.1; --

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_Like_cupin.
PFam; PF00190; Cupin; 2.
FT NON_TER 580
SQ SEQUENCE 580 AA; 66575 MW; 247AA144CCC1F36D CRC64;
Alignment Scores:
Pred. No.: 3,47e-181 Length: 580
Score: 2887.50 Matches: 562
Percent Similarity: 97.11% Conservative: 9
Best Local Similarity: 95.58% Mismatches: 6
Query Match: 81.06% Indels: 11
DB: 2 Gaps: 6
US-10-728-323-1 (1-2032) x Q6PSU3 (1-580)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGGATCTTGTCTGCTTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCATCTTACCAAGAAAGAAAACAGAGAACCCCTGCCCCAG 169
Db 21 SerAlaThrGlnAlaIys---SerProTyrr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGTGCCCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGCATGCCAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuIysGlnIysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTGTATGATCTCTCGAGGACACATCGC 289
Db 59 ArgCysThrIysLeuGluTyrrAspProArgCysValTyrrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGCGACACGTCGCCGCCCAACCCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyrr 94
QY 350 GATGATGACCCCGCTCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409
Db 95 AspAspAspArgGlnProArgGlnGluGlyArgGlyArgGlyArgGlyArgGlyArgGly 114
QY 410 AGGGAGCGGTGAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 469
Db 115 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY 470 CATCAGAGCCACCGAAAATTAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgIysIleArgProGluGlyArgGlyGluGlyGluGlnGluTrpGlyThr 154
QY 530 CCAGGTAGCCATGTGAGGAGAAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrrPheProSerArg 174
QY 590 CGGTTTAGCCCGCTACGGGAACCAAAACCGGTAGGATCCGGGTCTCTCAGAGGTTTCAAC 649
Db 175 ArgPheSerThrArgTyrrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTCAGATCGAGGCCAAA 709
Db 195 GlnArgSerIysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTCTGTGTTCTTCCCAAGCACGCTGATGCTATAACATCTTGTGTTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProIysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTGACCGTGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAACTTTGACG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCACTTCTTACATCTTGAACCGCCATGCACAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrrIleLeuAsnArgHisAspAsn 274


```
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 275 GlnAsnLeuArgValAlaIysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTCTTTCAGGCGCTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrluGlnGlyPheSerArgAsnThr 314
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGGAGGTGCTGTGTAGAAGAAAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
QY 1070 GCAGGAGGTGACCAAGAGGAGAGCGGCGAGCGGATGGAGTACTCGGAGTACTGAGAAC 1129
Db 335 AlaGlyGlyGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTGTAGTCAAGTGTCAAGGAGCAGCTTCAAGAACTTACTTAAGCACGCT 1189
Db 354 AsnGluGlyValIleValIleValSerLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTCTCAAGAAAGCTCCGAAGAGAGGAGATATACCAACCAATCAACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAGCGAGCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGACCCCGCTTCCAGGACCTGGACATCATCTCACCTGTGTAGAGATCAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTTGTATGTCCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGTCACAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAACCTGGAACCTTGAACCTGGCTGTGTAAAGAAAGAGCAACACAGAGGGACGGCG 1489
Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 472
QY 1490 GAA-----CAAGAGGAGGAGCAAGACAGCAAGAGAGGAGGAAGTAAACAGAGAGTG 1540
Db 473 GluGlnGluTrpGluGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTGAAAGGAGCGATGTGTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGCCATCAAGCTTCTCCGAACCTCCATCTCGTTGGCTTGGGTATCAACGCTGAAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCCAGAAATCTTCTTCAGGTGATAGGACAAATGTGATACACAGATAGAGAGCAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTTCCCTGGGTGGGTGAAACAAGTTGAGAAGCTCATCAAAAACAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCATTGTGAGTGCT 1804
Db 573 ArgGluSerHisPheValSerAla 580

RESULT 4
Q6PSU4 PRELIMINARY; PRT; 428 AA.
AC Q6PSU4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
```

OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581851; AAT00596.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
FT NON TER 1
SQ SEQUENCE 428 AA; 48094 MW; FLA32DAEC33AD06F CRC64;

Alignment Scores:
Pred. No.: 3,028-134 Length: 428
Score: 2172.50 Matches: 427
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 60.99% Indels: 1
DB: 2 Gaps: 1

US-10-728-323-1 (1-2032) x Q6PSU4 (1-428)

QY 644 TTTTGACCAAGGTCAAGCGAGTTTCAGAAATCTCCAGAATCACCGTATTGTGAGATCGAG 703
Db 2 PheAspGlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGlu 21
QY 704 GCCAAACCTTAACACTCTTGTCTTCTCCAGACGCTGATGCTGATAACATCTTGTATC 763
Db 22 AlaLysProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIle 41
QY 764 CAGCAAGGCAAGCCACCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 823
Db 42 GlnGlnGlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeu 61
QY 824 GACGAGGCGCATGCTCAGAAATCCCATCCGCTTTCATCTTCTTCTTCTTCTTCTTCT 883
Db 62 AspGluGlyHisAlaLeuArgIleProSerGlyPheIleSerTyrlleLeuAsnArgHis 81
QY 884 GACAAACCAAGACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTT 943
Db 82 AspAsnGlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPhe 101
QY 944 GAGGATTTCTCCCGGAGCAGCGAGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCT 1003
Db 102 GluAspPhePheProAlaSerSerArgAspGlnSerSerTyrlleGlnGlyPheSerArg 121
QY 1004 AATAGCTTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGCGTGTCTTGTAGAA 1063
Db 122 AsnThrLeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGlu 141
QY 1064 GAGAAATGCAAGGAGTGTAGTCAAGTGTCAAAGGACGAGCGATGAGTACTCGGAGTAGT 1123
Db 142 GluAsnAlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSer 161
QY 1124 GAGAAATGAAGGAGTGTAGTCAAGTGTCAAAGGACGAGCGTGTGAAGAACTTACTTAAG 1183
Db 162 GluAsnAsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLys 181
QY 1184 CACGCTAAATCCGTCTCAAGAAAGCGCTCCGAAGAGAGGAGATATACCAACCAATC 1243
Db 182 HisAlaLysSerValSerLysLysGlySer---GluGluGlyAspIleThrAsnProIle 200
QY 1244 AACTTGAGAAAGGCGAGCCCGATCTTTCTAAACAACCTTTGGGAAGTATTTTGGAGTGAAG 1303
Db 201 AsnLeuArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLys 220

QY	1304	CCAGACAAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATC	1363
Db	221	ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIle	240
QY	1364	AAAGAGGAGCTTGTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCTCGTCGTC	1423
Db	241	LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal	260
QY	1424	AACAAAGGAACCTCGAATCTCGTGGCTGTAAAGAAAAGAGCAACAAACAGAGGGGA	1483
Db	261	AsnLysGlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGly	280
QY	1484	CGCGCGAAGACAGAGGAGCAAGACGAAAGAGAGCGGAGTAACAGAGAGTGGCT	1543
Db	281	ArgArgGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArg	300
QY	1544	AGGTACACAGCCAGGTGAAGNAGCGGATGTTTCATCATGCCAGCAGCTCATCCAGTA	1603
Db	301	ArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProVal	320
QY	1604	GCATCAACCGCTTCTCCGAATCTCAATCTCGTTCGGCTTCGGTATCAACGCTCAAAACAAAC	1663
Db	321	AlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn	340
QY	1664	CACAGAACTTCTCTGAGGTGATAGGACAATGTGATAGACAGATAGAGCAAGCG	1723
Db	341	HisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAla	360
QY	1724	AAGGATTTAGCATCTCCCTGGTGGGTGAACAGTTGAGAAGCTCATCAAAACAGAG	1783
Db	361	LysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLys	380
QY	1784	GAATCTCACTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAG	1843
Db	381	GluSerHisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlu	400
QY	1844	AAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAAAACCAAGGAGGAGGTCACATC	1903
Db	401	LysGluSerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeu	420
QY	1904	CTTTCATTTTGAAGCTTTTAAC	1927
Db	421	LeuSerIleLeuLysAlaPheAsn	428
RESULT 5			
Q6PSUS			
ID	Q6PSUS	PRELIMINARY;	PRT; 299 AA.
AC	Q6PSUS;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Conarachin (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]_TaxID=3818;		
RP	SEQUENCE FROM N.A.		
RA	Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY581850; AAT00595.1; -.		
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin_region.		
DR	InterPro; IPR011051; RmC_like_cupin.		
DR	Pfam; PF00190; Cupin; 1.		
FT	NON TER		
SQ	SEQUENCE 299 AA; 33604 MW; 594C3AB0C2FD49EB CRC64;		
Alignment Scores:			
Pred. No.:	8.28e-90	Length:	299

RESULT 6

US-10-728-323-1 (1-2032) x Q6PSUS (1-299)			
QY	1010	TTGGAGGCGCCTTCAATGCGGAATTTCAATGAGATACGAGGGTGCTGTTAGAAGAGAT	1069
Db	1	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn	20
QY	1070	GCAGAGGTGAGCAAGAGAGAGAGGCGAGCGATCGAGTACTCGGAGTAGTGAGAAC	1129
Db	21	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	40
QY	1130	AATGAGGAGTGATAGTCAAGTGTCAAAGAGACGCTTGAAGAACTTACTAAGCACGCT	1189
Db	41	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	60
QY	1190	AAATCCGTCTCAAGAAAAGGCTCCGAAGAAGAGGAGGATATCACCAACCCCAATCACTG	1249
Db	61	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	80
QY	1250	AGAGAGCGAGCCCGATCTTTCTAAACAACTTTGGAAAGTTATTGAGGTGAAGCCAGAC	1309
Db	81	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	100
QY	1310	AAGAAGACCCCGAGCTTCAGACCTGCACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	101	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu	120
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCGTCAACAAA	1429
Db	121	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	140
QY	1430	CGAACTGGAACCTTGAATCTCGTGGCTGTAAAGAAAGACCAACACAGAGGAGCGCGG	1489
Db	141	GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArg	160
QY	1490	GAAGAGAGGAGGAGCAAGACGAGAGAGGAGGAGTAACAGAGGTCGCTAGGTAC	1549
Db	161	GluGluGluGluAspGluAspGlnGluGluGlySerAsnArgGluValArgArgTyr	180
QY	1550	ACAGCGAGTGTGAAGAGGCGATGTGTTTCATGTCAGGAGCTCATCCAGTAGCCATC	1609
Db	181	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	200
QY	1610	AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGATATCAACGCTGAAAAACACACAGA	1669
Db	201	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	220
QY	1670	ATCTTCTTCGAGGTGATAAGGACAAATGTGTAGACCAAGATAGAGACGACGGAAGAT	1729
Db	221	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	240
QY	1730	TTAGCATTTCCCTCGGTCGGGTGAACAAGTTGAGAGAGCTCATCAAAAAACAGAGAAATCT	1789
Db	241	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnArgGluSer	260
QY	1790	CACTTTGTGAGTGTGCTCCTCAATCTCAATCTCAATCTCCGTCGTCTCTCTGAGAAAGAG	1849
Db	261	HisPheValSerAlaArgProGlnSerGlnSerProSer	273
QY	1850	TCTCTGTAGAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGCTCCACTCTTTCA	1909
Db	274	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	293
QY	1910	ATTTTGAAGGCTTTTAAC	1927
Db	294	IleLeuLysAlaPheAsn	299

```

Q6PSU6
ID Q6PSU6 PRELIMINARY; PRT; 303 AA.
AC Q6PSU6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581849; AAT00594.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_like_cupin.
DR Pfam; PF00190; Cupin; 1.
FT NON TER 1
SQ SEQUENCE 303 AA; 34133 MW; 5B4E21BEB48654EC CRC64;

Alignment Scores:
Pred. No.: 7,84e-86 Length: 303
Score: 1435.00 Matches: 291
Percent Similarity: 95.50% Conservative: 6
Best Local Similarity: 93.57% Mismatches: 2
Query Match: 40.29% Indels: 12
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x Q6PSU6 (1-303)
QY 1004 AATACGTGGAGGCGCCCTTCATGCGGATTCATCAGATACGGAGGTGCTGTAGAA 1063
DB 2 AsnThrLeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGlu 21
QY 1064 GAGAACTCAGGAGGTGAGCAAGAGGAGAGAGGCGGATGAGTACTCCGAGTAGT 1123
DB 22 GluAsnAlaGlyGlyGluGlnGluGluGluGluGluGluGluGluGluGluGluGlu 41
QY 1124 GAGAACTAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1183
DB 42 Asp---AsnGluGlyValIleValIleValIleValIleValIleValIleValIle 60
QY 1184 CACGCTAAATCCCTCTCAAGAAAGGCTCCGAGAGAGGAGGAGATATCACCAACCAATC 1243
DB 61 HisAlaLysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIle 79
QY 1244 AACTTGAGAGAGGCGAGCCGATCTTCTAACAACTTTGGGAGATTATTGAGGTGAAG 1303
DB 80 AsnLeuArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLys 99
QY 1304 CCAGACAAGAAAGACCCCGAGCTTCAGGACCTCGACATGATGCTACCTGTGTGAGATC 1363
DB 100 ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIle 119
QY 1364 AAGAGAGGAGCTTTGATGCTCCACACATCAACTCAAGGCCATGTTATCGTCGTC 1423
DB 120 LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal 139
QY 1424 AACAAAGGAACCTGAAACCTTGAACTCGTCGCTGTAAGAAAGAGCAACAAAGAGGGGA 1483
DB 140 AsnLysGlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnGly 159
QY 1484 CGCGGGGAA-----GAAGAGGAGGACCAAGACCAAGAGAGGAGGAGGAGTAACAGA 1534
DB 160 ArgArgGluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGlu 179
QY 1535 GAGTCGCTAGGTACACAGCAGGAGGTGAAGAGGCGATGTTCATCATGCCAGCAGCT 1594

```

```

Db 180 GluValArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAla 199
QY 1595 CATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTCTCTGGTTCGGTATCAACGCT 1654
Db 200 HisProValAlaIleAsnAlaSerSerGluLeuHisLeuGlyPheGlyIleAsnAla 219
QY 1655 GAAACCAACACACAGATCTCTTCAGAGTGATAGGACAATGTGATACCAAGATAGAG 1714
Db 220 GluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGlu 239
QY 1715 AAGCAAGCGAAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAGTTTGAGAAGCTCATCAA 1774
Db 240 LysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLys 259
QY 1775 AACCAGAGGAATCTCATTCTGTGAGTGCTCTCTCAATCTCAATCTCAATCTCCGTCG 1834
Db 260 AsnGlnArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 277
QY 1835 TCTCCTGAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGAGGAGAAACCAAGAGGAGAG 1894
Db 278 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLys 292
QY 1895 GGTCCACTCTCTTCAATTTTGAAGGCTTTTAAC 1927
Db 293 GlyProLeuLeuSerIleLeuLysAlaPheAsn 303

RESULT 7
Q7XXT2 PRELIMINARY; PRT; 621 AA.
AC Q7XXT2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Prepro beta-conglycinin alpha prime subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Urade R., Nakatani H., Nakano C.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB113351; BAC78524.1; -.
DR PDB; 1UIK; X-ray; A/B/Cs-.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 621 AA; 72247 MW; 0DF4B6E33737B7F CRC64;

Alignment Scores:
Pred. No.: 5.1e-77 Length: 621
Score: 1302.00 Matches: 286
Percent Similarity: 61.02% Conservative: 121
Best Local Similarity: 42.88% Mismatches: 170
Query Match: 36.55% Indels: 90
DB: 2 Gaps: 19

US-10-728-323-1 (1-2032) x Q7XXT2 (1-621)
QY 50 ATCAGAGGAGGAGGTTTCTCCACTGATGCTGTGTGCTAGGATCTTGTCTGGCTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
QY 110 TCTGCAACCCATGCCCAAGTCATCCTTACACAGAGAAACACAGAACCCCTGCGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36
QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229

```

Db 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACC-----AAGCTCGAGTATGATCTCTCGTTGT-----GTCTATGAT 271
 Db 57 ArgCysAsnLeuLeuLysValGluGluGluGluGluCysGluGluGlyGlnIleProArg 76
 QY 272 CTTCTGA---GGACACACTGGCACCCACCAACCAACAGTTCCTCCCTCCAGGGGAGCGGACACGT 328
 Db 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyGluLysGluGluAspGlu 96
 QY 329 GCGCCGCAACCCGAGACTACGATGATGACCGC---CGTCAACCCCGAGAGAGGNA--- 382
 Db 97 GlyGluGlnProArgProPheProPheProArgGlnProArgGlnProArgGlnGluGlu 116
 QY 383 -----CGAGGCCGATCG 394
 Db 117 HisGluGlnLysGluGluHisGluTyrHisArgLysGluGluLysHisGlyGlyLysGly 136
 QY 395 GGA-----CCAGCTGGACCGGAGGAGCGGTGAA 421
 Db 137 SerGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156
 QY 422 AGAGAAGAA---GACTGGAGAACCAACAGAGAGATTGGAGCGGACCAAGTATCAGCAG 478
 Db 157 GluGluLysHisGluTyrGlnHisLysGlnGluLysHisGlnGlyLysGluSerGluGlu 176
 QY 479 CCACGGAAATAGGCCCGGAAGGAGAGAGAGAACAAAGAG---TGGGGAACACCCAGGT 535
 Db 177 GluGluGluAspGlnAspGluAspGluGluGlnAspLysGluSerGlnSerGluGly 196
 QY 536 AGCCATGTGAGGAGAGAAACATCTCGGAAC-----AACCTTTCTACTTCCCGTCA 586
 Db 197 SerGluSerGlnArgGluProArgArgHisLysAsnLysAsnProPheHisPheAsnSer 216
 QY 587 AGCGGTTTAGCCCGCTACGGGAACCAACGAGTAGGATCGGGTCTCGCAGAGTTT 646
 Db 217 LysArgPheGlnThrLeuLysAsnGlnTyrGlyHisValArgValLeuGlnArgPhe 236
 QY 647 GACCAAGGTCAAGGAGTTCAGAACTCCAGAACTCCAGATTCACCGTATTGTGCAGATCGAGGC 706
 Db 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgGluLeuGluPheAsnSer 256
 QY 707 AACCTTAACACTTCTTCTTCCAGCAGCGTGTATGATAACATCTCTGTTATCCAG 766
 Db 257 LysProAsnThrLeuLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValIleLeu 276
 QY 767 CAAAGGCAACCCGCTGACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGAC 826
 Db 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuGln 296
 QY 827 GAGGGCCATGCACTCAGAATCCCATCCGGTTTCATTCTCTACATCTTGAACCGCCATGAC 886
 Db 297 SerGlyAspAlaLeuArgValProAlaGlyThrTyrTyrValValAsnProAspAsn 316
 QY 887 AACCAAGACCTCAGAGTAGCTAAATCTCCATGCCGTTTAACACACCCCGCCAGTTTGAG 946
 Db 317 AspGluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGlu 336
 QY 947 GATTTCTTCCCGGAGCAGCGGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAAT 1006
 Db 337 SerPhePheLeuSerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerLysAsn 356
 QY 1007 ACCTTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGCGTGTGTAGAAGAG 1066
 Db 357 IleLeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArg 376
 QY 1067 AATGCAGAGGTGAGCAA---GAGGAGAGAGCGGCAGAGCGGATGCGAGTACTCGGAGTAGT 1123
 Db 377 GluGluGlyGlnGlnGlnGlyGluGluArgLeuGln----- 388
 QY 1124 GAGAACAATGAGGAGTGTAGTCAAGTGTCAAGGAGCAGCGTTGAGAACTTACTAAG 1183
 Db 389 -----GluSerValIleValGluIleSerLysLysGlnIleArgGluLeuSerLys 405

QY 1184 CACGCTAAATCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACCAATC 1243
 Db 406 HisAlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPhe 422
 QY 1244 AACTTGAGAGAAGGCGAGCCGATCTTCTTAACAACTTTGGGAAGTTATTTAGGTGAAG 1303
 Db 423 AsnLeuArgSerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIleThr 442
 QY 1304 CCAGACAAGAAAGACCCAGCTTCAGGACCTCGGACATGCTCACCTGTGTACAGATC 1363
 Db 443 ProGlu---LysAsnProGlnLeuArgAspLeuAspValPheLeuSerValValAspMet 461
 QY 1364 AAAGAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTCTCGTC 1423
 Db 462 AsnGluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIle 481
 QY 1424 AACAAAGAACTGGAAACCTTGAACCTCTGGCTGTAAAGAAAGACACAACAGAGGGA 1483
 Db 482 AsnGluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGln 500
 QY 1484 CGCGCGGAGAGAGGAGGAGGACGACGACGAGAGGAGGAGTAACAGAGAGGTGGT 1543
 Db 501 GlnGlnGluGluGlnPro-----LeuGluValArg 510
 QY 1544 AGGTACACAGCGAGGTTGAAGGAGCGGATGTTTCATCATGCCAGCAGCTCATCCAGTA 1603
 Db 511 LysTyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProVal 530
 QY 1604 GCCATCAACGTTCTCCGAACTCATCTCTGTTGGTTCGTTATCAACGCTGAAACAAAC 1663
 Db 531 ValValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsn 550
 QY 1664 CACAGAACTTCTTCGAGGTGATAAGCACAACTGTGTAGACCCAGATAGAGAGCAAGCG 1723
 Db 551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570
 QY 1724 AAGGATTTAGCATTCCTCGGTCGGGTGCAACAAAGTTGAGAAGCTCATCAAAACAGAG 1783
 Db 571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590
 QY 1784 GAATCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCCGCTCTCTGAG 1843
 Db 591 GluSerTyrPheValAspAlaGlnProGln----- 600
 QY 1844 AAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTC 1903
 Db 601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613
 QY 1904 CTTTCAATTTTGAAGGCTTTT 1924
 Db 614 SerSerIleLeuArgAlaPhe 620
 RESULT 8
 Q6EBC1 PRELIMINARY; PRT; 533 AA.
 ID Q6EBC1; Q6EBC1; 28, Created)
 AC 25-OCT-2004 (TREMELrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)
 DE Beta-conglutinin.
 OS Lupinus albus (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monteiro S.A., Freitas R.M., Teixeira A.N., Ferreira R.B.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY500372; AAS97865.1; -;
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.

DR InterPro: IPR007113; Cupin region.
DR InterPro: IPR011051; RmlC_like_cupin.
DR Pfam: PF00190; Cupin; 2.
SQ SEQUENCE 533 AA; 62129 MW; F17E49A6F2C4E9BC CRC64;

Alignment Scores:
Pred. No.: 3,02e-76 Length: 533
Score: 1290.00 Matches: 273
Percent Similarity: 62.90% Conservative: 117
Best Local Similarity: 44.03% Mismatches: 132
Query Match: 36.22% Indels: 98
DB: 2 Gaps: 17

US-10-728-323-1 (1-2032) x Q6BEC1 (1-533)

QY 44 GGAGCAATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGATCCTTGTCTGGCT 103
DB 2 GlyLysMetArgValArgPheProThrLeuValLeuValLeuGlyLeuValPheLeuMet 21
QY 104 TCAGTTCTCGCAACGCATGCCAAGTCATCCTTACCAGGAAGAACAGAGACCCCTGC 163
DB 22 AlaValSerIleGlyLeuAla-----TyrGlyLysAsp----- 33
QY 164 GCCCAGAGGTGCTCCAGAGTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATGC 223
DB 34 -----ValLeuLysSerHisGluArgProGluGluArgGluGluGln----- 47
QY 224 GAGTCTCGTCCACCAAGCTCGAGTATGATCTCTGTGTGTCTATGATCTCCAGGACAC 283
DB 48 -----GluTrpGlnProArg----- 52
QY 284 ACTGGCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCGGA 343
DB 53 -----ArgGlnArgPro----- 56
QY 344 GACTACGATGATGACCCGCTCAACCCGAGAGAGAGGAGGAGCGGATGGGACCAAGCT 403
DB 57 -----GlnSerArgArgGluGluArgGluGluGlnGluGluGlnGlnGlnGln 73
QY 404 GGAACGAGGAGCGGTAAAGAGAGAACTGGAGACCAACCAAGAGAAGATTCGAGCGGA 463
DB 74 TyrProArgArgGlnSerGlyTyrGlu-----ArgArg 84
QY 464 CCAAGTCATCAGACGCCACCGAAATAAGCCCGAGAGAGAGAGAGAGAGAGAGAGTGG 523
DB 85 GlnTyrHisGlu-----ArgSerGluGlnArgGluGluGluGluGlnGlnGln 102
QY 524 GGAACACCGAGTACGATGTGAGGAAGAAACATCTCGGAACACCTTTCTACTTCCCG 583
DB 103 GlySerPro---SerTyrSerArgArgGln-----ArgAsnProTyrHisPheSer 118
QY 584 TCAAGCGGTTTACGACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGG 643
DB 119 SerGlnArgPheGlnThrLeuTyrLysAsnArgAsnGlyLysIleArgValLeuGluArg 138
QY 644 TTTGACCAAGGTCACAGGAGTTTCAGAACTCCAGAACTCAGGATCCCGTATTCGAGATCG 703
DB 139 PheAspGlnArgThrAsnArgLeuGluAsnLeuGlnAsnTyrArgIleValGluPheGln 158
QY 704 GCCAAACCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATC 763
DB 159 SerLysProAsnThrLeuIleLeuPheLysHisSerAspAlaAspTyrValLeuValVal 178
QY 764 CAGCAAGGCGCAACCGGTAGCCGTAGCAAAATGGCAATAAACAGAGAGCTTTAATCTT 823
DB 179 LeuAsnGlyArgAlaThrIleThrIleValAsnProAspArgGlnAlaTyrAsnLeu 198
QY 824 GACGAGGCGCATGCATCAGAAATCCCATCCGGTTTCATTCCTACATCTTGAACCCGCAT 893
DB 199 GlutryGlyAspAlaLeuArgIleProAlaGlySerThrSerTyrIleLeuAsnProAsp 218
QY 884 GACACCCAGAACCTCAGAGTAGCTAAATCTCCATCCCGTTTAAACACCCCGCCAGTTT 943

219 AspAsnGlnLysLeuArgValValLysLeuAlaIleProIleAsnAsnProGlyTyrPhe 238
944 GAGGATTTCTCCCGCGAGCAGCCGAGACCAATATCTCTCTTTCAGGCGCTTTCAGCAGG 1003
239 TyrAspPheTyrProSerSerThrLysAspGlnGlnSerTyrPheSerGlyPheSerArg 258
1004 AATACGTTTCGAGCGCCCTTCAATCGGAATCAATGAGATACGAGGAGGTCGTGTTA--- 1060
259 AsnThrLeuGluAlaThrPheAsnThrArgTyrGluGluIleGlnArgIleIleLeuGly 278
1061 ---GAAGAGAATGTCAGGAGGTGAGCAAGAGAGAGGCGAGGCGATGGAGTACTCGG 1117
279 AsnGluAspGluGlnGluTyrGluGluGlnArgArgGlyGlnGluGln----- 294
1118 AGTAGTGAAACATGAAAGAGGTAGTCAAGTGTCAAAGAGTGTCAAAGAGCAGCTTGAAGACTT 1177
295 ---SerAspGlnAspGluGlyValIleValIleValSerLysLysGlnIleGlnLysLeu 313
1178 ACTAAGCAGCTAAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGGAGGATATCACAAC 1237
314 ThrLysHisAlaGlnSerSerSer-----GlyLysAspLysProSerAsp---SerGly 330
1238 CCAATCAACTTGAGAGAGCGAGCCCGATCTTTCTAAACAACCTTTGGGAGTATTATTGAG 1297
331 ProPheAsnLeuArgSerAsnGluProIleTyrSerAsnLysTyrGlyAsnPheTyrGlu 350
1298 GTGAAGCCAGACAGAAAGAACCCCGAGCTTCCAGGACCTCGACATGATGCTACCTGTGTA 1357
351 IleThrProAsp---ArgAsnProGlnValGlnAspLeuAsnIleSerLeuThrTyrIle 369
1358 GAGATCAAGAGAGGAGCTTGTGATGCTCCACACTTCAACTCAAGGCCATGGTATTCGTC 1417
370 LysIleAsnGluGlyAlaLeuLeuLeuProHisTyrAsnSerLysAlaIleTyrValVal 389
1418 GTCGTCAACAAAGGAACCTTGAACCTGCTGCTGTAAAGAAAGAGCAACCAACAG 1477
390 ValValAspGluGlyGluGlyAsnTyrGluLeuValGlyIleArgAspGlnArgGln 409
1478 AGGGGACGGCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1537
410 -----GlnAspGluGlnGluGluGlu-----GluGlu 419
1538 GTGCGTAGTACACAGCGAGTTGAAGAGGCGGTGCTTCATCATGCGCAGCGCTCAT 1597
420 ValIleArgTyrSerAlaArgLeuSerGluGlyAspIlePheValIleProAlaGlyTyr 439
1598 CCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAA 1657
440 ProlSerIleAsnAlaSerSerAsnLeuArgLeuLeuGlyPheGlyIleAsnAlaAsp 459
1658 AACCAACACAGAAATCTTCTTCGAGGTATAGGACAATGTGTATAGACAGATAGAGAAG 1717
460 GluAsnGlnArgAsnPheLeuAlaGlySerLysAspAsnValIleArgGlnLeuAspArg 479
1718 CAAGCGAAGATTTAGCATTCCTCGGTCGGGTGAGCAAGTTGAGAGCTCATCAAAAC 1777
480 AlaValAsnGluLeuThrPheProGlySerAlaGluAspIleGluArgLeuIleLysAsn 499
1778 CAGAAGGAATCTCAGTCTTGTGAGTGTCTCTCTCAATCTCAATCTCTCAATCTCTCGCTCT 1837
500 GlnGlnGlnSerTyrPheAlaAsnGlyGlnProGlnGlnGlnGlnGln----- 515
1838 CCTGAGAAAGATCTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1897
516 -----GlnGlnSerGluLysGluGlyArgArgGly 525

RESULT 9
Q948Y0 PRELIMINARY; PRT; 621 AA.
ID Q948Y0
AC Q948Y0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Beta-conglycinin alpha prime subunit.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Fukazawa C.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB030838; BAB64303.1; -;
 DR HSSP; P25974; 11PU.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR007113; Cupin region.

DR InterPro; IPR011051; RmlC_Like_cupin.

DR Pfam; PF00190; Cupin; 2.

SQ SEQUENCE 621 AA; 72138 MW; 0196CAD30CB566DC CRC64;

Alignment Scores:

Pred. No.:	7,76e-76	Length:	621
Score:	1284.00	Matches:	284
Percent Similarity:	59.56%	Conservative:	118
Best Local Similarity:	42.07%	Mismatches:	167
Query Match:	36.05%	Indels:	106
DB:	2	Gaps:	19

US-10-728-323-1 (1-2032) x Q948Y0 (1-621)

QY 50 ATGAGAGGAGGTTTCTCCATGATGCTGTGTCTAGGATCCTGTCTGCTTCAGTT 109
 DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAAACAGAGAACCCCTGCCCCAG 169
 DB 20 SerValSerPheClyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36
 QY 170 AGTGCTCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCATCGCAGTCT 229
 DB 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGTGTGACC-----AAGTCGAGTATGATCTCTGTGTGT-----GTCTATGAT 271
 DB 57 ArgCysAsnLeuLeuLysValGluGluGluGluGluCysGluGluGluGlnIleProArg 76
 QY 272 CTTGCA---GGACACTGGCACCCACCAACCAACGTTTCCCTCCAGGGAGCGACAGT 328
 DB 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyGluLysGluGluAspGlu 96
 QY 329 GGCCTGCAACCCGAGACTACGATGATGACCCG---CGTCAACCCCGAAGAGAGGAAGGA 385
 DB 97 GlyGluGlnProArgProPheProPheProArgProArgGlnProArgGlnGlu----- 114
 QY 386 GGCCTGATGGGACCACTGCGACGAGGAGCGTGAAGAGAGAAAGACTGGAGACAACA 445
 DB 115 -----GlyGluHisGluGlnLysGluGluHisGluGluHisGluTrpHisArgLys 128
 QY 446 AGAGAA-----GATTGGAGCGACCAAGTCATCATCAGCAGCCACGG 484
 DB 129 GluGluLysHisGlyGlyLysGlySerGluGluGlnAspGlyArgGluHisProArg 148
 QY 485 AAAATAAGGCCCC----- 496
 DB 149 ProHisGlnProHisGlnLysGluGluLysHisGluTrpGlnHisLysGlnLys 168
 QY 497 -----GAAGGAAGAGAGGAACAAGAG----- 520
 DB 169 HisGlnGlyLysGluSerGluGluGluGluGluAspGlnAspGluAspGluGlnAsp 188
 QY 521 -----TGGGGAACACCAAGTACCATGTGAGGGAAGAAACATCTCGGAAC----- 565
 DB 189 LysGluSerGlnGluSerGluGlySerGluSerGlnArgGluProArgArgHisLysAsn 208

QY 566 ---AAACCTTTCTACTTCCCTCAGGCGGTTTAGCACCCGCTACCGGAACCAAAACCGT 622
 DB 209 LysAsnProPheHisPheAsnSerLysArgPheGlnThrLeuPheLysAsnGlnTyrGly 228
 QY 623 AGGATCCGGGCTCTGCAGAGGTTTGCACCAAGAGTCAAGGCGAGTTTCAGAAATCCAGAA 682
 DB 229 HisValArgValLeuGlnArgPheAsnLysArgSerGlnGlnLeuGlnAsnLeuArgAsp 248
 QY 683 CACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGTCTTCCCAAGCACCTCAT 742
 DB 249 TyrArgGlnLeuGluPheAsnSerLysProAsnThrLeuLeuProHisAlaAsp 268
 QY 743 GCTGATAACATCTTGTATTATCCAGCAAGCAAGCCACCGTCACCTAGTACCAATGCCAAT 802
 DB 269 AlaAspTyrLeuIleValIleLeuAsnGlyThrAlaIleLeuThrLeuValAsnAsnAsp 288
 QY 803 AACAGAAAGAGCTTTAATCTTGCAGAGGCGCATGCTCAGAAATCCCATCCGTTTCATT 862
 DB 289 AspArgAspSerTyrAsnLeuGlnSerGlyAspAlaLeuArgValProAlaGlyThrThr 308
 QY 863 TCCTACATCTTGAACCGCCATGCACCAACCAACCTCAGAGTAGCTAAATCTCCATGCC 922
 DB 309 TyrTyrValValAsnProAspAsnAspGluAsnLeuArgMetIleThrLeuAlaIlePro 328
 QY 923 GTTAACACACCCGGCCAGTTTGCAGGATTTCTTCCCGCGAGCAGCGCAGACCAATCATCC 982
 DB 329 ValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGlnAlaGlnGlnSer 348
 QY 983 TACTTGACGGGTTTCAGCAGGAATACGTTGAGGCGCGCTTCAATCGGAATTCATCAG 1042
 DB 349 TyrLeuGlnGlyPheSerLysAsnIleLeuGluAlaSerTyrAspThrLysPheGluGlu 368
 QY 1043 ATACGAGGCGTCTGTTAGAGAGAAATCAGAGGTGAGCAA---GAGCAGAGAGGGCAG 1099
 DB 369 IleAsnLysValLeuPheGlyArgGluGluGlnGlnGlnGlyGluGluArgLeuGln 388
 QY 1100 AGCGATGAGTACTCGGAGTAGTGAGAAACAATGAAGAGTAGTAGTCAAAAGTGTCAAAG 1159
 DB 389 -----GluSerValIleValGluLysSerLys 397
 QY 1160 GAGCAGGTTGAAGAACTTACTAAGCAGCTAAATCCGCTCAAGAAAGAGGTCGCGAAGA 1219
 DB 398 LysGlnIleArgGluLeuSerLysArgAlaLysSerSerSerArgLysThrIleSerSer 417
 QY 1220 GAGGAGATATCACCAACCCCAATCACTTGAGAGAGCGAGCCGATCTTCTTAACAAC 1279
 DB 418 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 434
 QY 1280 TTTGGAAAGTTATTGTAGGTGAAGCCAGACAAGAAAGAACCCCGCTTCAGGACCTGGAC 1339
 DB 435 LeuGlyLysLeuPheGluIleThrProGlu---LysAsnProGlnLeuArgAspLeuAsp 453
 QY 1340 ATGATGTCACCTGTGTAGATCAAGAAGAGCTTTGTGCTCCCGACACTTCACTCA 1399
 DB 454 ValPheLeuSerValValAspMetAsnGluGlyAlaLeuPheLeuProHisPheAsnSer 473
 QY 1400 AAGGCATGTTATCGTCTGCTCAACAAAGAACTGGAAACCTTGAACCTCGTCTCTA 1459
 DB 474 LysAlaIleValValLeuValIleAsnGluGlyGluAlaAsnIleGluLeuValGlyIle 493
 QY 1460 AGAAAGAACCAACACAGAGGGGACGCGGGAAGAGAGGAGGACGAAGACGAAGAGAG 1519
 DB 494 ---LysGluGlnGlnArgGlnGlnGlnGluGlnPro----- 506
 QY 1520 GAGGGAAGTAAACAGAGAGGTCGCTAGGTACACAGGAGGTTGAAGAGAGGCGATGTTC 1579
 DB 507 -----LeuGluValArgLysTyrArgAlaGluLeuSerGluGlnAspIlePhe 522
 QY 1580 ATCATGCCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACCTCATCTGTTGGC 1639
 DB 523 VallieProAlaGlyTyrProValValValAsnAlaThrSerAspLeuAsnPhePheAla 542
 QY 1640 TTCGATATCAACGCTGAAACCAACACAGAAATCTTCTTCGAGGTGATAAGACAAATGTG 1699

```

Db      543 PheGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlySerIlyAspAsnVal 562
QY      1700 ATAGACAGATAGAGAGCAAGCGAAGATTTAGCATTCCTCGGTGGGTGAACAGTT 1759
Db      563 IleserGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaIlyAspIle 582
QY      1760 GACAAGCTCATCAAAACAGAGAGGATCTCACTTTGTGAGTCTCGTCTCAATCTCAA 1819
Db      583 GluAsnLeuIlySerGlnSerGluSerTyrPheValAspAlaGlnProGln----- 600
QY      1820 TCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCTCGAGAAAGAGGATCAAGAGGAGGAA 1879
Db      601 -----GlnIlyGluGluGly 605
QY      1880 AACCAAGAGGAGGAGGTCACCTCTTCAATTTTGAAGCTTTT 1924
Db      606 AsnIlyGlyArgIlyGlyProLeuSerSerIleLeuArgAlaIle 620

RESULT 10
Q94LX2 PRELIMINARY; PRT; 605 AA.
AC Q94LX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-conglycinin alpha subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21327318; PubMed=11434464;
RA Yoshino M., Kanazawa A., Tsutsumi K., Nakamura I., Shinamoto Y.;
RT "Structure and characterization of the gene encoding alpha subunit of
RT soybean beta-conglycinin.";
RL Genes Genet. Syst. 76:99-105(2001).
DR EMBL; AB051865; BAB56161.1; -.
DR PIR; S20007; S20007.
DR HSSP; P25974; 1IPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin_region.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 605 AA; 70306 MW; 8ACB6F8532662984 CRC64;

Alignment Scores:
Pred. No.: 1,04e-75 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.37% Conservative: 108
Best Local Similarity: 43.05% Mismatches: 173
Query Match: 35.99% Indels: 96
DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q94LX2 (1-605)
QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTGTCTGCTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuGluGlyValPheLeuAlaSerVal 19
QY 110 TTGCAACGATGCCAAGTACACCTTACCAGAGAAACAGAGAACCCCTGCGCCAC 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProIlyHisAsn 36
QY 170 AGGTGCTCCAGGTTGTCAACAGGACCGGATGATCTGAAGCAAAAGGATCGCGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCCAGTATGAT-----CCTCGTTGT 262

```

```

Db      57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyLeuIleProArg--- 75
QY      263 GTCTATGATCTCTCGAGGACACACTGGCACCAACCAACAGTTTCCCTCCAGGGAGCGG 322
Db      76 -----ProArgProArg-----ProGlnHisProGlu 84
QY      323 ACACGTGGCGCGCAACCCGAGACTACGATGATGAC----- 358
Db      85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluGlnProArgProIle 104
QY      359 -----CGCGCTCAACCCCGAAGAGAG-----GAAAGAGGCCGA 391
Db      105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY      392 TGG-----GGACAGCTGACCGGAGGAGCGTGAAGAGAGAA 430
Db      125 TrpProArgLysGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144
QY      431 GACTGGAGACACCAAGAGAGAT-----TGGAGGCGACCAAGTCAATCAGCAGCCACGG 484
Db      145 AspGluGlnAspGluArgGlnPheProPheProArgProProHisGlnIlySerGlu 164
QY      485 AAATAAGAGCCCGAAGAGAGAGAGAAAGAGTGGGGAACACACAGGTAGCCATGTG 544
Db      165 ArgLysGlnGluGluAspGluAspGluGlnGlnArgGluSerGluSerGluAsp 184
QY      545 AGGAA-----GAAACATCTCGGAACCAACCTTTCTCTCCGTCAAGCGGTTTAGC 598
Db      185 SerGluLeuArgArgHisIlyAsnIlyAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY      599 ACCCGCTACGGGAACCAAAACGTTAGGATCCGGTCTCGAGAGTTTACCAAAAGGTCA 658
Db      205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY      659 AGCGAGTTTCAATCTCCAGATCTCCAGATCCGCTATGTGTCAGATCGAGCGCAAACTAACCT 718
Db      225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerIlySerProAla 244
QY      719 CTGTGTTCTTCCCAAGCAGCTGATGCTGATACATCTCTGTTTATCCAGCAAGGCAAGCC 778
Db      245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY      779 ACCGTGACCGTAGCAATGCGCAATACCAAGAGAGCTTTAATCTTGACGAGGCGCATGCA 838
Db      265 IleLeuSerLeuValAsnAsnAspAspAspSerTyrArgLeuGlnSerGlyAspAla 284
QY      839 CTCAGATCCCATCCGCTTTCATTTCTACATCTTGAACCGCATGCAACACAGAACCTC 898
Db      285 LeuArgValProSerGlyThrThrTyrValValAsnProAspAsnAsnGlnAsnLeu 304
QY      899 AGAGTAGCTAAATCTCCATGCCCTTAACACACCCGCGCAGTTTGAGATTCTTCCCG 958
Db      305 ArgLeuIleThrLeuAlaIleProValAsnIlySerProGlyArgPheGluSerPhePheLeu 324
QY      959 GCGACGACCGGAGACCAATCATCTCTGAGGCTTCAGCGAGTTCAGCAAGTACGTTGAGGCC 1018
Db      325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY      1019 GCCTCAATTCGAATTCATGAGATACGAGGCTGTGTTAGAGAGAATGCAGAGGT 1078
Db      345 SerTyrAspThrLysPheGluGluIleAsnIlyValLeuPheSerArgGluGluGly--- 363
QY      1079 GAGCAAGAGGAGAGGAGGAGGAGGATGAGTACTCGGAGTAGTGAGAACATGAAGA 1138
Db      364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
QY      1139 GTGATAGTCAAGTGTCAAGGAGCACCTTGAAGAACTTACTAAGCAGCTAAATCCGTC 1198
Db      375 ValIleValGluIleSerLysGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY      1199 TCAAGAAAGGCTCCGAGAGAGGAGGAGATATCACCAACCCCAATCACTTGAAGAGGC 1258
Db      395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411

```



```
QY 1259 GAGCCCGGATCTTTCTTAACAACCTTTGGGAAGTTATTTGAGTGAGCGCAGACGAAGAAGAAC 1318
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 412 AppProLileTySerAsnLysLeuGlyLysPhePheGluLeuThrProGlu---LysAsn 430
QY 1319 CCCCAGCTTCAGGACCTGACATGATGCTCACTGTGTGATGATCAAAAGAGAGCTTTG 1378
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCACACTTCAACTCAAGGCGATGTTATCGTCTGTCGTCACAAAGAACTGGA 1438
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 451 LeuLeuProHiePheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCTGCTGTAGAAAGAGCAACACAGAGCGGCGCGGAAAGAGAG 1498
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGCAAGAGAGGAGGAAGTAACAGAGAGGTGCTGATACACAGCGAGG 1558
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAGGCGATGTTTCATCATGCGCAGAGCTCATCCAGTACCCATCAACGCTTCC 1618
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY 1619 TCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCCTT 1678
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGACAAATGTATGATACCAAGATAGAGAAGCAAGCGAGGATTTAGCATTC 1738
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGlnLeuAlaPhe 559
QY 1739 CTGGTGGTGGTCAACAAGCTTGAGAGCTCATCAAAACACAGAGAAATCTCACATTGTTG 1798
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAG 1858
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGAGGAGAAACCAAGAGGAGGAGGTCATCTCTTTCAATTTTGAAG 1918
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 603 AlaPhe 604

RESULT 11
GLCA_SOYBN
ID _GLCA_SOYBN STANDARD; PRT; 605 AA.
AC P13916;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-conglycinin, alpha chain precursor.
OS Glycine max (Soybean)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=91355860; PubMed=2103438;
RA Sebastiani F.L., Farrell L.B., Schuler M.A., Beachy R.N.;
RT "Complete sequence of a cDNA of alpha subunit of soybean beta-
conglycinin."
RL Plant Mol. Biol. 15:197-201(1990).
CC -!- FUNCTION: Seed storage protein. Accumulates during seed
CC development and is hydrolyzed after germination to provide a
CC carbon and nitrogen source for the developing seedling.
```

```
CC -!- SUBUNIT: The alpha-, alpha-, and beta-subunits associate in
CC various combinations to form trimeric proteins.
CC -!- SUBCELLULAR LOCATION: Embryo axis, and cotyledonary membrane-bound
CC vacuolar protein bodies.
CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17698; CAA35691.1; -.
CC PIR; S14681; FWSYBA.
CC HSP; P25974; 1IFU.
CC InterPro; IPR006045; Cupin.
CC InterPro; IPR007113; Cupin region.
CC InterPro; IPR011031; Multihem cyt.
CC InterPro; IPR011051; RmlC_like_cupin.
CC Pfam; PF00190; Cupin; 2.
CC Glycoprotein; Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 22
FT PROPEP 23 62
FT CHAIN 63 605 Beta-conglycinin, alpha chain.
FT CARBOHYD 261 261 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 517 517 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 605 AA; 70293 MW; CBEBA30506BBEC57 CRC64;

Alignment Scores:
Pred. No.: 1,21e-75 Length: 605
Score: 1281.00 Matches: 285
Percent Similarity: 59.21% Conservative: 107
Best Local Similarity: 43.05% Mismatches: 174
Query Match: 35.96% Indels: 96
DB: 1 Gaps: 18

US-10-728-323-1 (1-2032) x GLCA_SOYBN (1-605)
QY 50 ATGAGAGGAGGAGTTCCTCCACTGATGCTGTGTAGGATCCTTGTCTGCTTCAGTT 109
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCGATGCCAAGTCATCCTTACCAGAAAGAAAACAGAGAACCCCTGCGCCACG 169
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGTGTCCTCCAGAGTTGTCAACAGAACCGGATGATCTTGAAGCAAAAGCATGCGAGTCT 229
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGTGCACCAAGCTCGAGTATGAT-----CCTGTTGT 262
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 57 ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyGluIleProArg--- 75
QY 263 GTCTATGATCTCGAGGACACATGCGCACCAACCAACCTTCCCTCCAGGGGAGCGG 322
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACAGTGGCCGCCCAACCCGAGACTTACGATGATGAC----- 358
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 85 ArgGluProGlnGlnProGlyGluLysGluAspGluAspGluGlnProArgProIle 104
QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAAGGAGCGCA 391
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGlnGlu 124
QY 392 TGG-----GGACCACTGACCGCGAGCGCGTGAAGAGAGAA 430
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 125 TrpProArgLysGluGluLysArgGlyGlyLysGlySerGluGluGluAspGlu 144
QY 431 GACTGAGACCAACCAAGAGAGAT-----TGGAGGCGACCAAGTATCATGAGCCACGG 484
```


Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 13

ID Q948X9 PRELIMINARY; PRT; 623 AA.
 AC Q948X9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Beta-conglycinin alpha-subunit.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Fukazawa C.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB030839; BAB64304.1; -;
 DR HSSP; P25974; IIPJ.

DR DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR011051; RmlC_Like_cupin.
 DR Pfam; PF00190; Cupin; 2.
 SQ - SEQUENCE 623 AA; 72475 MW; EBF033DA1316AE9A CRC64;

Alignment Scores:

Pred. No.: 1,28e-74 Length: 623
 Score: 1265.50 Matches: 283
 Percent Similarity: 58.22% Conservative: 110
 Best Local Similarity: 41.93% Mismatches: 179
 Query Match: 35.53% Indels: 103
 DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q948X9 (1-623)

QY 50 ATGAGAGGAGGTTTCTCCATGATGCTGTGCTAGGATCTTGTCTGCTGCTCAGTT 109
 Db 1 MetArgAlaAspPhePro-----LeuLeuLeuGlyValValPheLeuAlaSerVal 18
 QY 110 TCTGCAACGATGCCAAGTCATCATCTTACCAGAAACAGAGAACCCCTCGGCCAG 169
 Db 19 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProLysHisAsn 35
 QY 170 AGTGCTCTCAGAGTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCGATCGCAGTCT 229
 Db 36 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 55
 QY 230 CGCTGCACC-----AAGTCGAGTATGATCTCTGTTGTTGTTATGATCTCGAGGACAC 283
 Db 56 ArgCysAsnLeuLeuLysValGluLysGluGluCys-----GluGluGlyGlu 72
 QY 284 ACTGGCACCACCAACCAAGCTTCCCTCCAGGGGAGGACAGCTGGCGCCCAACCCGGA 343
 Db 73 IleProArgProArgProArgProGlnHisProGluArg---GluProGlnGlnProGly 91
 QY 344 GACTACGATGATGACCGC-----CGTCAA 367
 Db 92 GluLysGluGluAspGluAspGluGlnProArgProIleProPheProArgProArgGln 111
 QY 368 CCCCAGAGAGGAA----- 382
 Db 112 ProArgGlnGluGluGluHisGluGlnArgGluGluGlnGlnTrpProArgLysGluGlu 131
 QY 383 ---GGAGGCGATGGGA-----CCAGCT 403
 Db 132 LysArgGlyGlyLysGlySerGluGluGlnAspGlyArgGluHisProArgProHis 151
 QY 404 GGACCCGAGGAGCTGCAAGAGAGAAAGACTGGAGACCAACCAAGAGAGTGGAGCGCA 463
 Db 152 GlnProHisAspGluAspGluGluGlnAspGluArgGln-----PheProPheProArg 169

QY 464 CCAAGTCATCAGCACCCACCGGAAATAAGG-----CCCGAA 499
 Db 170 ProProHisGlnLysGluSerGluArgLysGlnGluAspGluAspGluGln 189
 QY 500 GGAAGAGAGGAGAAACAAGAGTGGGAAACACACAGGTAGCATGTGAGGGAAGAACATCT 559
 Db 190 GlnArgGluSerGluGluSerGluSerGluSerGlnArgGluLeuArgHisLys 209
 QY 560 CGAAACAACCTTTCTACTCCCGTCAAGCGGTTTAGCACCCGTTACCGGAAACCAAAAC 619
 Db 210 AsnLysAsnProPheHisPheGlySerAsnArgPheGluThrLeuPheLysAsnGlnTyr 229
 QY 620 GGTAGATCCGGTCTCGCAGAGTTTACCAAAAGGTCAAGGCAGTTTTCAGATCTCCAG 679
 Db 230 GlyArgIleArgValLeuGlnArgPheAsnGlnArgSerProGlnLeuGlnAsnLeuArg 249
 QY 680 AATCACCGTATTGTGCAGATCGAGGCCAAACCTAAACACTCTTGTCTTCCCAAGCACGCT 739
 Db 250 AspTyrArgIleLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProAsnHisAla 269
 QY 740 GATGCTGATAACATCCTTGTATTATCAGCAAGGGCAGCCCGTACCGTACCGTAATGTC 799
 Db 270 AspAlaAspTyrLeuIleAlaIleLeuAsnGlyThrAlaIleLeuSerLeuValAsnAsn 289
 QY 800 AATAACAGAAAGCTTTAATCTTCAGCGGCCCATGCATCTCAGATCCCATCCGTTTC 859
 Db 290 AspAspArgAspSerTyrArgLeuGlnSerGlyAspAlaLeuArgValProSerGlyThr 309
 QY 860 ATTTCTCTACATCTTCAACCGCATCAGCAACCAACCTCAGAGTAGTAAATCTCCATG 919
 Db 310 ThrTyrTyrValValAsnProAsnAsnGluAsnLeuArgLeuIleThrLeuAlaIle 329
 QY 920 CCCGTTAAACACACCGCGCAGTTTTCAGGATTTCTTCCCGCGCAGCAGCGACCAATCA 979
 Db 330 ProValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGluAlaGlnGln 349
 QY 980 TCCTACTTGCAGGCTTTCAGCAGGAATACGTTGGAGGCGCTTCAATCGGAAATTCAT 1039
 Db 350 SerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAlaSerTyrAspThrLysPheGlu 369
 QY 1040 GAGATACGAGGCTGCTGTAGAGAAGATGACGAGGTGAGCAAGAGAGAGAGAGAGCAG 1099
 Db 370 GluIleAsnLysValLeuPheSerArgGluGly-----GlnGlnGlnGlyGlu 386
 QY 1100 AGCGCATCGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTGATAGTAAAGTGCAAG 1159
 Db 387 GlnArgLeu-----GlnGluSerValIleValGluIleSerLys 399
 QY 1160 GAGCAGTGAAGAACTTACTAAGCACCGCTAATCGTCTCAAGAAAGGCTCCGAAAGAA 1219
 Db 400 GluGlnIleArgAlaLeuSerLysArgAlaLysSerSerSerArgLysThrIleSerSer 419
 QY 1220 GAGGAGATATCACCACCACTCACTTGAGAGAGGCGGAGCGCATCTTCTTAACAAC 1279
 Db 420 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 436
 QY 1280 TTTGGGAAGTATTTGAGGTGAAGCAGACAAAGAAACCCCGCTTCAGGACCTGGAC 1339
 Db 437 LeuGlyLysPhePheGluIleThrProGlu---LysAsnProGlnLeuArgAspLeuAsp 455
 QY 1340 ATGATGTCACCTGTGTAGAGATCAAGAGAGGCTTTGATGCTCCACACTTCACTCA 1399
 Db 456 IlePheLeuSerIleValAspMetAsnGluGlyAlaLeuLeuLeuProHisPheAsnSer 475
 QY 1400 AAGGCGCATGTTATCGTCTGCTCAACAAAGAACTGGAAACCTTGAACCTCGGTGTA 1459
 Db 476 LysAlaIleValIleLeuValIleAsnGluGlyAspAlaAsnIleGluLeuValGlyLeu 495
 QY 1460 AGAAAGAGCAACAACAGAGGGGACCGCGGGAAGAGAGGAGGAGCAAGACGAAGAG 1519
 Db 496 ---LysGluGlnGln-----GluGluGlnGln 504

```

QY 1520 GAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGGAAGAAAGCGATGTGTC 1579
Db      |||
505 GluGluGlnProLeuGluValArgLysTyrArgAlaGluLeuSerGluGlnAspIlePhe 524
QY 1580 ATCATGCCAGCAGCTCATCCAGTAGCCATCAAGCTTCCTCCGAACTCCATCTGCTGGC 1639
Db      :::::|||||
525 ValIleProAlaGlyTyrProValValValAsnAlaThrSerAsnLeuAsnPheAla 544
QY 1640 TTCGGTATCAACGGCTGAAACCAACACACATCTCTTCCTGACGGTGATAAGGACAATGTG 1699
Db      |||
545 IleGlyIleAsnAlaGlnAsnGlnArgAsnPheLeuAlaGlySerGlnAspAsnVal 564
QY 1700 ATAGACCATAGAGAACGAAAGGATTAGCATTCCTCGGTGCGGTGAACAAAGTT 1759
Db      |||
565 IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaGlnAlaVal 584
QY 1760 GAGAACTCATCAAAACCAAGAAAGTCTCACTTTGTGAGTCTCGTCTCAATCTCAA 1819
Db      |||
585 GluLysLeuLeuLysAsnGlnArgGluSerTyrPheValAspAlaGlnProLysLys--- 603
QY 1820 TCTCAATCTCCGTCTCTCGAGAAAGAGTCTCTCGAGAAAGAGATCAAGAGGAGGAA 1879
Db      604 -----LysGluGluGly 607

QY 1880 AACCAAGGAGGAAGGTCCATCTCTTCAATTTGAAGGCTTTT 1924
Db      |||
608 AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe 622

RESULT 14
Q41674 PRELIMINARY; PRT; 545 AA.
ID Q41674;
AC Q41674;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conviclin precursor.
OS Vicia marbonensis (Narbonne vetch).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosoid I; Fabales; Fabaceae; Papilionoideae; Vicieae; Vicia.
OX NCBI_TaxID:3912;
RN [1]_TaxID:3912;
RC SEQUENCE FROM N.A.
RP - TISSUE=Cotyledon;
RA Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 271986; CA896513.1; -.
DR HSP; P25974; IIPD.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_Iike_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 545 AA; 62810 MW; 459A876F92F5A87E CRC64;

Alignment Scores:
Pred. No.: 3,56e-74 Length: 545
Score: 1258.50 Matches: 279
Percent Similarity: 58.18% Conservative: 98
Best Local Similarity: 43.06% Mismatches: 150
Query Match: 35.33% Indels: 121
DB: 2 Gaps: 16

US-10-728-323-1 (1-2032) x Q41674 (1-545)
QY 38 GPAGCAGGACCAATGAGAGGGAGGTTCCTCCACTGATGCTGTGCTAGGGATCCTGTGC 97
Db      |||
1 MetAlaThrAlaMetLysSerArgPhe---ProValLeuLeuLeuGlyIleIle 19
QY 98 CTGGCTTCAGTTTCTGCAACGCATGCCAAGTCATCACCTTACCAGAAACAGAGAAC 157

```

```

Db      |||
20 LeuAlaSerLeuCysValThrTyrAlaAsn-----TyrAspGluGlyThrGlu--- 35
QY 158 CCCTCGCGCCAGAGGTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAG 217
Db      :::::|||||
35 ----- 35
QY 218 GCATCGAGTCTCGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTGTCTATGATCCTCGA 277
Db      |||
36 -----ProArgVal----- 38
QY 278 GGACACATCGGCACCAACCAACCAAGTTCCTCCAGGGAGCGACACAGTGGCGGCCAA 337
Db      |||
39 -----ProGlyGlnArgGluArgGlyArgGln 47
QY 338 CCCGAGACTACGATGATGACCGCGTCAACCCCGAAGAGGAGGAGCGGATGGGA 397
Db      |||
48 GluGlyGluLysGluGluLysArgHis-----GlyLutrpArg 60
QY 398 CCA-----GCTGGACCGAGGAGCGTGAAGA----- 424
Db      |||
61 ProSerHisGluLysGluAlaGlnProGlyArgArgGluArgTrpGluThrSerGluGlu 80
QY 425 -----GAAGAAGACTGGAGA-----CAACCAAGAGAAGATTGGAGGCGACA 466
Db      |||
81 GluGluArgValAspGluLutrpArgGlySerGlnArgHisGluAsp-----Pro 97
QY 467 AGTCATCAGCAGCCAGGAAATAAGCCCAAGCAAGAGAGGAGAAACAAGAGTGG--- 523
Db      |||
98 GluGluArgAlaArgGluArgLysArgAlaGluGluArgGluArgArgGlnTrpGlu 117
QY 524 GGAACACCCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCG 583
Db      |||
118 GluGluLysGluLysSerSerLysSerGlnGluArgArgAsnProPheLeuPheLys 137
QY 584 TCAAGCGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGAGTCCGGGTCTGCAGAGG 643
Db      |||
138 SerAsnLysPheLeuThrLeuPheGluAsnGluAsnGlyHisIleArgArgLeuGlnArg 157
QY 644 TTTGACCAAAAGTCAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAG 703
Db      |||
158 PheAspLysArgSerAspLeuPheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArg 177
QY 704 GCCAAACCTAACACTCTTTGTTCTTCCCAAGCAGCGTGATGCTGATAACATCTTGTATC 763
Db      |||
178 AlaLysProHisThrIlePheLeuProGlnHisIleAspAlaAspLeuIleLeuThrVal 197
QY 764 CAGCAAGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAACAGAAAGCTTTAATCTT 823
Db      |||
198 LeuSerGlyArgAlaIleLeuThrValLeuSerProAsnAspArgAsnSerTyrAsnLeu 217
QY 824 GACGAGGCCCATGCACTCAGAAATCCCATCCGTTTTCATTTCCTACATCTTGAACCGCAT 883
Db      |||
218 GluArgGlyAspThrIleLeuSerProAlaGlyThrThrSerTyrLeuLeuAsnGlnAsp 237
QY 884 GACAACCAAGAACCTCAGAGTAGCTAAATCTCCATGCGCGCTTAACACACCCGCGCAGTTT 943
Db      |||
238 AspGluGluAspLeuArgValAlaAspLeuSerIleSerValAsnArgProGlyLysVal 257
QY 944 GAGGATTTCTCCCGGCGAGCAGCGGAGACCAACATCATCTTACTTGCAGGGCTTCAGCAGG 1003
Db      |||
258 GluSerPheGlyLeuSerGlySerLysAsnGln-----TyrLeuArgGlyPheSerLys 275
QY 1004 AATAGTTTGGAGGCGCTTCAATCGGAATTCATAGATACGAGGGGTCTGTGTAGAA 1063
Db      |||
276 AsnIleLeuGluAlaSerLeuAsnThrLysTyrGluThrIleGluLysValLeuLeuGlu 295
QY 1064 GAGAAATGCAAGGAGTGAGCAAGAGGAGAGGCGAGCGGATGAGTACTCGGAGTAGT 1123
Db      |||
296 Glu-----ProGlnGlnSerIleGlyGlnLysArgArgSerGlnArgGlnGlu 311
QY 1124 GAGAAACAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGGTGGAAGAACTTACTAAG 1183
Db      |||

```

Db 312 ThrAsn-----AlaLeuVallySerArgGluGlnValGluGluLeuLysArg 328
QY 1184 CACGCTAAATCCCTCTCAAGAAGCGTCGGAAGAAGAGGAGATATACCAACCCCAATC 1243
Db LeuAlaLysSerSerLysLysGlyValSerSerGluPheGlu-----ProPhe 345
QY 1244 AACCTCAGAGAAGCGGAGCGGATCTTCTTAACAACCTTTGGGAAGTTATTGAGGTGAAG 1303
Db AsnLeuArgSerGlnAsnProLysTyrSerAsnLysPheGlyLysLeuPheGluLeuThr 365
QY 1304 CCAGAAGAAGAACCCCGAGCTTCAGGACCTCGACATGATGCTACCTGTGTAGATGATC 1363
Db ProGluLysLysTyrProGlnLeuGlnAspLeuAspLeuPheValSerSerValGluLeu 385
QY 1364 AAAGAAGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGCTGCTC 1423
Db AsnGluGlyGlyLeuMetLeuProHisTyrAsnSerArgAlaLeuValleLeuLeuVal 405
QY 1424 AACAAAGGAAGCTGAAACCTTGAACTCGTGCTGTAAAGAAAGAGCAACAAAGAGGGGA 1483
Db AsnGluGlyLysGlyAsnLeuGluLeuValGlyLeuLysAsnGlnGlnGlnArg 425
QY 1484 CGCGCGGAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGTGCGT 1543
Db GluArgGluAspGluGln-----GlnValGln 434
QY 1544 AGGTACACAGCGAGGTGAAGGAGCGATGTTTCATCATGCGACAGCTCATCCAGTA 1603
Db ArgTyrGluAlaArgLeuSerProGlyAspValValleLeuProAlaGlyHisProVal 454
QY 1604 GCATCAACGCTTCCCGAATCCATCTCGTTCGCTTCGTTATCAACCTGAAACAAAC 1663
Db AlaValSerAlaSerAsnLeuLeuLeuGlyPheGlyLeuAsnAlaGluLeuAsn 474
QY 1664 CACAGATCTTCTTCGAGGTATAAGGACAAATGTGATACACAGATAGAGAAGCAAGCG 1723
Db GlnArgAsnPheLeuThrGlySerAspAspAsnValleSerGlnleLeuAsnProVal 494
QY 1724 AAGATTTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAGTCAATCAAAAACCAAGAG 1783
Db LysGluLeuThrPheProGlySerAlaGlnGluValAsnArgLeuLeuLysAsnGlnGlu 514
QY 1784 GAATCTCACTTTGTAGTGTCTCTCTCAATCTCAATCTCAATCTCGTCTCTCTGAG 1843
Db HisSerHisPheAlaAsnAlaGluPro----- 523
QY 1844 AAAGAGTCTCTCAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1903
Db 524 -----GluGlnLysGlyGluGluSerGlnArgLysArgSerProIle 537
QY 1904 CTTTCAATTTTGAAGCTTTTAAC 1927
Db SerSerIleLeuGlyThrPheAsn 545
RESULT 15
Q9M3X6 PRELIMINARY; PRT; 613 AA.
ID AC Q9M3X6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Convicilin precursor.
GN Names=cvc;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed cotyledon;
RA Casey R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ276875; CAB82855.1; --
DR HSP; P25974; IIPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 30 613 convicilin.
SQ SEQUENCE 613 AA; 72063 MW; 49AED99F6135DD19 CRC64;
Alignment Scores:
Pred. No.: 1-148-73 Length: 613
Score: 1251.00 Matches: 283
Percent Similarity: 57.61% Conservative: 107
Best Local Similarity: 41.80% Mismatches: 155
Query Match: 35.12% Indels: 132
DB: 2 Gaps: 17
US-10-728-323-1 (1-2032) x Q9M3X6 (1-613)
QY 68 CCACATGATCTGCTAGGATCCTGTCTGCTGCTTCA---GTTTCTCAACGATGCC 124
Db 10 ProLeuLeuLeuLeuGlyIlePheLeuAlaSerValValCysValThrTyrAla 29
QY 125 AAGTCATCATCTTACAGAGAAGAAACAGAG---AACCCCTGCGCCAGAGGTGCTCCAG 181
Db 30 Asn-----TyrAspGluGlySerGluProArgValProAlaGlnArg-----Glu 44
QY 182 AGTTGTCAACAGACCCGATGACTTGAAGAAAAGGATCGAGTCTCGTGCACCAAG 241
Db 45 ArgGlyArgGlnGlnGlyGlyLysGluGlyLysArgHisGlyGluTrpArgProSer--- 63
QY 242 CTCGATGATGATCTCTGCTGCTGCTATGATCTCTCGAGGACACACTGGCACCAACCA 301
Db 64 -----TyrGlu-----LysGluGlu 68
QY 302 CTTTCCCTTCAGGGAGCGGACAGTGGCCGCCCAACCCGAGAGATGATGATGATGCCG 361
Db 69 AspGluGluGluGlyGlnArgGluArgGlyArgGlnGluGlyGlyGlyGlyGlyArg 88
QY 362 CGTCAACCCGAG 403
Db 89 His-----GlyGluTrpArgProSerTyrGlnLysGlnGluAsp 101
QY 403 ----- 403
Db 102 GluGluGluLysGlnLysTyrArgTyrGlnArgGluLysGluAspGluGluLysGln 121
QY 404 -----GGACCGAGGAG 415
Db 122 LysTyrGlnTyrGlnArgGluLysLysGluGlnLysGluValGlnProGlyArgGluArg 141
QY 416 CGTGAAG 460
Db 142 TrpGluArgGluGluAspGluGlnValAspGluGluTrpArgGlySerGlnArgArg 161
QY 461 -----CGACCAAGTCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 493
Db 162 GluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLysArgAspArg 181
QY 494 CCCGAAGGAAG 553
Db 182 ArgHisGlnArgGlyGlyGluGluGlu-----GluArgSerSerGluSer 196
QY 554 ACATCTCGGAACAACCTTTCTACTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 613
Db 197 GlnGluArgArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeuPheGluAsn 216
QY 614 CAAAACGGTAGGATCCGGGTCTCGAGAGTTTACCAAGAGGTTCAGAGGAGGAGGAGGAG 673
Db 217 GluAsnGlyHisIleArgLeuLeuGlnArgPheAspLysArgSerAspLeuPheGluAsn 236

```
QY 674 CTCAGAATCACCGTATTGTGCAGATCGAGGCCAAACCTAAACACTCTTGTCTTCTCCCAAG 733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 LeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePheLeuProGln 256
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 CAGCGTGTGTCGATACACTCCTTGTATCCAGCAAGGGCAAGCCACCGTAGCCGTAGCA 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 HisIleAspAlaAspLeuIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeu 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 794 AATGGCAATAACAGAGAGCTTTAATCTTGACGAGGGCCATGCATCTCAGATCCCATCC 853
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 SerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLysLeuProAla 296
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 854 GGTTCATTCTCTACACTTTGAACCCCATGACCAACAGACCTCAGAGTAGTAAATC 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 GlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgLeuValAspLeu 316
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 914 TCCATCCCGGTTAAACACCGGCCAGTTGTAGGATTCTCCCGCGAGCAGCCGAGAC 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 ValIleProValAsnGlyProGlyLysPheGluAlaPheAspLeuAlaLysAsnLysAsn 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 974 CAATCATCTACTTCAGGGCTTCAGCAGGAATACGTTGAGGCGCCCTTCAATCGGAA 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 Gln-----TyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerTyrAsnThrArg 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1034 TTCAATCAGATACGGAGGCTGCTGTGTAGAGAGAATGCAGGAGGTGAGCAAGAGGAGAGA 1093
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TyrGluThrIleGluLysValLeuLeuGluGlu-----GlnGluLys 368
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1094 GGGCAGAGCGGATGGAGTACTCGGAGTAGTGAGACAATGAAGGAGTGATAGTCAAGTG 1153
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 AspArgLysArg-----ArgGlnGlnGlyGluGluThrAspAlaIleValLysVal 385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1154 TCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCGTAAATCCGTCTCAAAGAAAGGCTCC 1213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 SerArgGluGlnIleGluGluLeuLysLysLeuAlaLysSerSerLysLysSerLeu 405
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1214 GAAGAGAGGGAGATATCAACCACTTGAAGAGAGGGCGGAGCCCGATCTTTCT 1273
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 ProSerGluPheGlu-----ProIleAsnLeuArgSerHisLysProGluTyrSer 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1274 AACAACTTTGGAGTATTTCAGGTGAAGCCAGACAGAAAGAACCCCGCTTCAGGAC 1333
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 AsnLysPheGlyLysLeuPheGluIleThrProGluLysLysTyrProGlnLeuGlnAsp 442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1334 CTGGACATCATGCTACCTGTGTAGAGATCAAAGAGGAGCTTGTATGCTCCACACTTC 1393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 LeuAspLeuPheValSerCysValGluIleAsnGluGlyAlaLeuMetLeuProHisTyr 462
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1394 AACTCAAGCCATGGTTATCGTCGTCTCAACAAAGGAACCTTGAACCTCGTG 1453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 463 AsnSerArgAlaIleValValLeuLeuValAsnGluGlyLysGlyAsnLeuGluLeuLeu 482
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1454 GCTGTAAGAAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGAGGACCAAGACGAA 1513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GlyLeuLysAsnGlnGlnGlu-----GluAspArg 494
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1514 GAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGTACACAGCGAGGTGAAGGAGCGCAT 1573
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 LysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSerProGlyAsp 512
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1574 GTGTTTCATCATCCAGCAGCTCATCCAGTAGCCATCAACGGTTCTCTCCGAACTCCATCTG 1633
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 ValValIleIleProAlaGlyHisProValAlaIleThrAlaSerSerAsnLeuAsnLeu 532
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1634 CTTGGCTTCGGTATCAACGCTGAAAACACACAGAACTTCTCTGAGGTGATAAGGAC 1693
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 LeuGlyPheGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeuSerGlySerAspAsp 552
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1694 AATGTGATAGACAGAGAGCAAGCGAAGGATTTAGCATTCCTCGGTCGGGTGAA 1753
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 AsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGlySerValGln 572
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1754 CAAGTTGAGAAGCTCATCAAAAAACAGAAAGGAATCTCACTTTGTGAGTGCTCGTCTCAA 1813
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 GluIleAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaAsnAlaGluPro--- 591
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1814 TCTCAATCTCAATCTCCGTCGTCCTCCTGAGAAAGAGTCTCCTGAGAAAGAGATCAAGAG 1873
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 -----GluGlnLysGlu 595
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1874 GAGGAAAACCAAGGAGGAGGTCCTCCTTTCAATTTGAAGGCTTTT 1924
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 GlnGlySerGlnGlyLysArgSerProLeuSerSerIleLeuGlyThrPhe 612
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 24, 2005, 10:04:41
Job time : 157.355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 09:03:06 ; Search time 35.1903 Seconds
(without alignments)
8620.958 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3562
Sequence: 1 aataacatatattcatc.....cgttgtggtgtttcttc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat 23082005 124356 29238/app query.fasta_1.4757
-DB=Issued Patents AA -OFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1.46 -runat 23082005 124356 29238 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3278	92.0	626	4	US-09-106-872A-4
2	1700	47.7	335	4	US-09-106-872A-17
3	1279	35.9	605	1	US-07-955-905A-24
4	1267	35.6	571	1	US-07-955-905A-25
5	1068	30.0	410	1	US-07-955-905A-26
6	811	22.8	566	1	US-07-955-905A-2
7	811	22.8	566	1	US-07-955-905A-22
8	788	22.1	421	1	US-07-955-905A-27
9	777	21.8	587	1	US-07-955-905A-23
10	591.5	16.6	489	4	US-09-424-283-3
11	569	16.0	524	4	US-09-424-283-1
12	559.5	15.7	448	4	US-09-323-195A-18

13	514.5	14.4	409	4	US-09-424-283-4	Sequence 4, Appli
14	513.5	14.4	523	4	US-09-323-195A-17	Sequence 17, Appl
15	501	14.1	444	4	US-09-424-283-2	Sequence 2, Appli
16	183	5.1	762	4	US-09-252-991A-29423	Sequence 29423, A
17	173	4.9	1564	4	US-10-144-198-2	Sequence 2, Appli
18	173	4.9	1564	4	US-10-144-198-4	Sequence 4, Appli
19	166.5	4.7	720	4	US-09-252-991A-21881	Sequence 21881, A
20	164.5	4.6	1266	4	US-09-252-991A-30851	Sequence 30851, A
21	163	4.6	714	4	US-09-949-016-10296	Sequence 10296, A
22	160	4.5	630	4	US-09-252-991A-26324	Sequence 26324, A
23	160	4.5	801	4	US-09-252-991A-29274	Sequence 29274, A
24	160	4.5	1898	1	US-08-056-200-94	Sequence 94, Appl
25	160	4.5	1898	2	US-08-800-644-94	Sequence 94, Appl
26	160	4.5	1898	4	US-09-538-092-1280	Sequence 1280, Ap
27	159	4.5	1043	4	US-09-252-991A-28885	Sequence 28885, A
28	158.5	4.4	1075	4	US-09-252-991A-18387	Sequence 18387, A
29	157	4.4	1780	1	US-08-769-309A-5	Sequence 5, Appli
30	157	4.4	1780	3	US-08-994-570-5	Sequence 5, Appli
31	157	4.4	1781	4	US-09-961-403-13	Sequence 13, Appl
32	157	4.4	2079	4	US-09-949-016-8301	Sequence 8301, Ap
33	155	4.4	754	4	US-09-976-594-375	Sequence 375, App
34	155	4.4	760	4	US-09-949-016-11129	Sequence 11129, A
35	155	4.4	863	4	US-09-252-991A-26099	Sequence 26099, A
36	154	4.3	650	4	US-09-252-991A-20375	Sequence 20375, A
37	153.5	4.3	1233	4	US-09-688-1888-89	Sequence 89, Appl
38	153.5	4.3	1233	4	US-09-291-417D-89	Sequence 89, Appl
39	152.5	4.3	580	4	US-09-248-796A-14473	Sequence 14473, A
40	152	4.3	1162	2	US-08-728-323A-2	Sequence 2, Appli
41	152	4.3	1162	3	US-09-298-568-2	Sequence 2, Appli
42	152	4.3	1162	4	US-09-410-399-2	Sequence 2, Appli
43	152	4.3	1162	4	US-09-894-273-2	Sequence 2, Appli
44	152	4.3	1427	4	US-09-252-991A-27005	Sequence 27005, A
45	151.5	4.3	917	4	US-09-252-991A-25101	Sequence 25101, A

ALIGNMENTS

RESULT 1

US-09-106-872A-4

; Sequence 4, Application US/09106872A

; Patent No. 6486311

; GENERAL INFORMATION:

; APPLICANT: Burks Jr., A. Wesley

; APPLICANT: Stanley, J. Steven

; APPLICANT: Cockrell, Gael

; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.

; APPLICANT: Helm, Ricki M.

; APPLICANT: Bannion, Gary A.

; TITLE OF INVENTION: Peanut Allergens and Methods

; FILE REFERENCE: HS 103 CIP

; CURRENT APPLICATION NUMBER: US/09/106,872A

; PRIOR FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: PCT/US96/15222

; PRIOR FILING DATE: 1996-09-23

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Arachis hypogaea

; FEATURE:

; OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 1

; OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 2

; OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 3

; OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 4

; OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 5

; OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,

QY 1550 ACACGAGGTGAAGGAGCGATGCTGTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCCCTCCGAATCCCATCTGCTTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCCTTCGAGGTATAGGCAATGTGATAGACAGATAGAGAGCAAGCGAAGGAT 1729
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTCGGTGAACAGTGTGAGAGCTCATCAAAACACAGAGGATCT 1789
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTTGAGTGTCTCTCAATCTCAATCTCAATCTCCGTCGTCTCTCGAGAAAGAG 1849
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCTCTTTCA 1909
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTGAGGCTTTTAAAC 1927
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 621 IleLeuLysAlaPheAsn 626

RESULT 2

US-09-106-872A-17

; Sequence 17, Application US/09106872A

; Patent No. 6486311

; GENERAL INFORMATION:

; APPLICANT: Burks Jr., A. Wesley

; APPLICANT: Stanley, J. Steven

; APPLICANT: Cockrell, Gael

; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.

; APPLICANT: Helm, Ricki M.

; APPLICANT: Bannon, Gary A.

; TITLE OF INVENTION: Peanut Allergens and Methods

; FILE REFERENCE: HS 103 CIP

; CURRENT APPLICATION NUMBER: US/09/106,872A

; CURRENT FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: PCT/US96/15222

; PRIOR FILING DATE: 1996-09-23

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17

; LENGTH: 335

; TYPE: PRT

; ORGANISM: Arachis hypogaea

US-09-106-872A-17

Alignment Scores:

Pred. No.: 8,73e-156 Length: 335

Score: 1700.00 Matches: 335

Percent Similarity: 99.41% Conservative: 0

Best Local Similarity: 99.41% Mismatches: 0

Query Match: 47.73% Indels: 2

DB: 4 Gaps: 1

US-10-728-323-1 (1-2032) x US-09-106-872A-17 (1-335)

QY 917 ATGCCGTTAAACACCCGCCAGTTTGGAGATTCTTCCCGGCGAGCAGCCAGACCAA 976
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 MetProValAsnThrProGlyGlnPheGluAspPheProAlaSerSerArgAspGln 20
 QY 977 TCATCTTACTTCAGGCGCTTCAGCAGGAATACGTGAGGCGCCCTTCAATGCGGAATTC 1036
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 21 SerSerTyrLeuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPheAsnAlaGluPhe 40

QY 1037 AATGATACGGAGGCTGCTTTAGAGAGAAATGCAGGAGCTCAGCAAGAGGAGAGGG 1096
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 41 AsnGluIleArgArgValLeuLeuGluAsnAlaGlyGlyGluGlnGluGluArgGly 60
 QY 1097 CAGAGCGGATGAGTAGTACTCGGAGTACTGAGAAACAATCAAGAGGAGTGTAGTCAAAAGTCTCA 1156
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIleValLysValSer 80
 QY 1157 AGGAGCAGCTTGAAGAACTTACTAAGCAGCGCTAAATCCGTCTCAAGAAAGGCTCCGAA 1216
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 81 LysGluHisValGluGluThrLysHisAlaLysSerValSerLysLysGlySerGlu 100
 QY 1217 GAAGGCGGAGATATCACCAACCAATCAACTGAGAGAGGCGGAGCCCGATCTTTCTAAC 1276
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 101 GluGluGlyAspIleThrAsnProIleAsnLeuArgGluGluProAspLeuSerAsn 120
 QY 1277 AACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCCCGAGCTTTCAGGACCTG 1336
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 121 AsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGlnLeuGlnAspLeu 140
 QY 1337 GACATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCACACTTCAAC 1396
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 141 AspMetMetLeuThrCysValGluIleLysGluGlyAlaLeuMetLeuProHisPheAsn 160
 QY 1397 TCAAGGCGCATGGTTATCGTCGTCAACAAGGAACCTGGAAACCTTGAACCTCGTGGCT 1456
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 161 SerLysAlaMetValIleValValAsnLysGlyThrGlyAsnLeuGluLeuValAla 180
 QY 1457 GTAAGAAAAGAGCAACAACAGAGGGGACCGCGGGAAGAGAGAGGACCAAGACGAAGAA 1516
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 181 ValArgLysGluGlnGlnGlnArgGlyArgGluGluGluGluAspGluGlu 200
 QY 1517 GAGGAGGAAGTAACAGAGAGGTGCTAGGTACACAGCAGGTTGAAGGAAGCGGATGTG 1576
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 201 GluGluGlySerAsnArgGluValArgArgTyrThrAlaArgLeuLysGluGlyAspVal 220
 QY 1577 TTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTT 1636
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 221 PheIleMetProAlaAlaHisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeu 240
 QY 1637 GCTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTCGAGGTGATAGGACAAT 1696
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 241 GlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsn 260
 QY 1697 GTGATAGACAGATAGAGAAAGCAAGAGGATTTAGCATTCCTCGGTCGGGTGAACAA 1756
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 261 ValIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGln 280
 QY 1757 GTTGAGAGCTCATCAAAACCAAGAGGAAATCTCACTTTGTGAGTCTCGTCTCTCAATCT 1816
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 281 ValGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAla-----GlnSer 298
 QY 1817 CAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTCGAGAAAGAGGATCAAGAGGAG 1876
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 299 GlnSerGlnSerProSerSerProGluLysGluSerProGluLysGluAspGlnGluGlu 318
 QY 1877 GAAAAACAGAGGAGGAGGTCACCTCTTTTCAATTTTGAAGGCTTTTAAAC 1927
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 319 GluAsnGlnGlyGlyLysGlyProLeuLeuSerIleLeuLysAlaPheAsn 335

RESULT 3

US-07-955-905A-24

; Sequence 24, Application US/07955905A

; Patent No. 5770433

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND

; TITLE OF INVENTION: PRECURSOR

; NUMBER OF SEQUENCES: 28

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS


```

Db      540 AlaGlySerGlnAspAenValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
Qy      1739 CCTGGGTCGGGTGAACAGTTGAGAGCTCATCAAAACAGAGGAATCTCACTTTGTG 1798
Db      560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAenGlnArgGluSerTyrPheVal 579
Qy      1799 AGTGCTCGTCTCAATCTCAATCTCCGTCGTCCTCCTGAGAAAGTCTCCTGAG 1858
Db      580 AspAlaGlnProLysLys----- 585
Qy      1859 AAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGTCCTCTTCAATTTTGAAG 1918
Db      586 -----LysGluGluGlyAenLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
Qy      1919 GCTTTT 1924
Db      603 AlaPhe 604

RESULT 4
US-07-955-905A-25
; Sequence 25, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pisum sativum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..571
; OTHER INFORMATION: /note= "Convicilin from P. sativum"
US-07-955-905A-25

Alignment Scores:
Pred. No.: 1..13e-113 Length: 571
Score: 1267.00 Matches: 276
Percent Similarity: 56.67% Conservative: 98
Best Local Similarity: 41.82% Mismatches: 146
Query Match: 35.57% Indels: 140
DB: 1 Gaps: 13

US-10-728-323-1 (1-2032) x US-07-955-905A-25 (1-571)
Qy      68 CCACTGATGCTGTGTAGGATCCTTGTCTCGCTTCAGTTTCTGCAAGCATGCCAAG 127
Db      10 ProLeuLeuLeuPheLeuGlyIleIlePheLeuAlaSerValThrTyrAlaAen 29
Qy      128 TCATCACCCTTACCAGAAGAAACAGAGAACCCCTCGCCCGCAGAGTGTCTCCAGAGTTGT 187
Db      29 ----- 29
Qy      188 CAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGTCTCGCTGCACCAAGCTCGAG 247
Db      29 ----- 29

```

```

Qy      248 TATGATCCTCGTGTGTCTATGATCCTCGAGACACACTGCGACCAACCAACCAACGTTCC 307
Db      30 -----TyrAspGluGlySerGluThrArgVal----- 38
Qy      308 CCTCAGGAGGAGCGGACACGTCGCGCCCAACCGGAGACTAGCATGATCAGCCGCGCTCAA 367
Db      39 ---ProGlyGlnArgGluArgGlyArgGlnGluGlyGluLysGluLysArgHis--- 56
Qy      368 CCCGGAAGAGAGGAGGCGCATGGGACACGATGGGACCGATGGGACCGAGGAGGAGCGT 418
Db      57 -----GlyGluTrpArgProSerTyrGluLysGluLysGluGlu 70
Qy      418 ----- 418
Db      71 GluLysGlnLysTyrArgTyrGlnArgGluLysLysGluGlnLysGluValGlnProGly 90
Qy      419 -----GAAAGAGAAGAACTGGAGACACCAAGAGAGATTGGAGG----- 460
Db      91 ArgGluArgTrpGluArgGluGluAspGluGlnValGluGluGluTrpArgGlySer 110
Qy      461 -----CGACCAAGTCATCAGCAGCCACCG 484
Db      111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLys 130
Qy      485 AAATAAGGCCCGAAGAGAGAGAGAGAACAAAGAGTGGGGAACACAGGTAGCCATGTG 544
Db      131 ArgAspArgArgHisGlnArgGluGlyGluGlu-----GluArgSer 145
Qy      545 AGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCTCAAGCGGCTTTAGCACCCGC 604
Db      146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165
Qy      605 TACGGGAACAAACCGTAGTAGTCCGGTCTCTGAGAGGTTTGACAAAGGTCAAGCGAG 664
Db      166 PheGluAsnGluAsnGlyHisIleArgArgLeuGlnArgPheAspLysArgSerAspLeu 185
Qy      665 TTTTCAGAACTCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTGT 724
Db      186 PheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205
Qy      725 CTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAAGGGCAAGCCACCGTG 784
Db      206 LeuProGlnHisIleAspAlaAspLeuLeuValLeuAsnGlyLysAlaIleLeu 225
Qy      785 ACCGTAGCAAAATGCGCAATAACAGAAAGACTTTAATCTTTCAGAGGCGCATGCACCTAGA 844
Db      226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245
Qy      845 ATCCCATCCGGTTTCATTTCTTCTACATCTTGAACCGCCATGACAAACAGAACCTCAGAGTA 904
Db      246 IleProAlaGlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
Qy      905 GCTAAATCTCCATGCCCGTTAACACACCGCGCCAGTTTGAGGATTTCTCCGCGGAGC 964
Db      266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
Qy      965 AGCCGAGACCAATCATCTTCTTTCAGGCGCTTTCAGCAGGAATACGTTGAGGCGCGCTTC 1024
Db      284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerLeu 303
Qy      1025 AATGCGGAATTCATGAGATACGGAGGTGCTTTAGAGAGAAATGCAGAGGTGAGCA 1084
Db      304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGlnLysLysProGln 323
Qy      1085 GAGGAGAGGCGCAGAGGCGATGAGTACTCCGAGTAGTGAGAAACAATCAAGAGGTGATA 1144
Db      324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
Qy      1145 GTCAAAGTGTCAAAGGAGCAGCTTGAAGAACTTCTTAAGCAGCGTAAATCTCTTCAAAG 1204
Db      341 IleLysValSerArgGluGlnIleGluLeuArgLysLeuAlaLysSerSerSerLys 360
Qy      1205 AAGGCTCGAAGAGAGGAGATATCACCAACCCATCATCTTGAGAGAGGCGAGCCC 1264

```

```

Db      361  LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
QY      1265  GATCTTTCTAAACACTTTGGGAGTTATTGTAGTGAACCCAGACAGAACAGACCCCCAG 1324
Db      378  GluTyrSerAsnLysPheGlyLysPheGluPheGluThrProGluLysLysTyrProGln 397
QY      1325  CTTACGAGCTGTGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTGATGCTC 1384
Db      398  LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
QY      1385  CCACACTTCAACTCAAGGCGCATGGTTATCGTCGTCTGCTCAACAAAGAACTCGAAACCTT 1444
Db      418  ProHisTyrAsnSerArgAlaIleValLeuValLeuValAsnGluGlyLysGlyAsnLeu 437
QY      1445  GAATCTGCTGCTTAAGAAAGAGCAACACAGAGGGGAGCGCGGAAAGAGGAGGAC 1504
Db      438  GluLeuLeuGlyLeuLysAsnGluGlnGlnGluArg----- 449
QY      1505  GAAGACGAAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAG 1564
Db      450  GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
QY      1565  GAAGGCGATGTGTTTCATCATGCGAGCAGCTATCCAGTAGCCATCAACGCTTCTCCGAA 1624
Db      468  ProGlyAspValIleIleProAlaGlyHisProValAlaIleSerAlaSerSerAsn 487
QY      1625  CTCCTACTCTGGCTTCGTTATCATACGCTGAACGCTGAACCAACACAGATCTTCTTCAGT 1684
Db      488  LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
QY      1685  GATAAGCAACATGTGTATACACAGATAGAGAGAGCAAGCAAGATTTAGCATTCCTCGG 1744
Db      508  SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
QY      1745  TCGGTGAACAAGTTGAGAGCTCATCAAAACCAAGAAAGATCTCACTTTCTGTAGTGTCT 1804
Db      528  SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysSerHisPheAlaSerAla 547
QY      1805  CGTCTCAATCTCAATCTCAATCTCGTCTCTCGTCTCTGAGAAAGAGTCTCTGAGAAAG 1864
Db      548  GluPro----- 550
QY      1865  GATCAAGAGGAGGAAACCAAGAGGAGGAGGTGCCACTCTTCAATTTTGAAGGCTTTT 1924
Db      551  GlnLysGluGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 5
US-07-955-905A-26
; Sequence 26, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pisum sativum
;
```

```

; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-410
; OTHER INFORMATION: /note= "vicilin from P. sativum"
US-07-955-905A-26

Alignment Scores:
Pred. No.: 1,77e-94 Length: 410
Score: 1068.00 Matches: 209
Percent Similarity: 71.64% Conservative: 84
Best Local Similarity: 51.10% Mismatches: 104
Query Match: 29.98% Indels: 12
DB: 1 Gaps: 5

US-10-728-323-1 (1-2032) x US-07-955-905A-26 (1-410)
QY      542  GTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTCCGCTCAAGCGGGTTTAGCACC 601
Db      13  ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32
QY      602  CGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAGGTCAAGG 661
Db      33  LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52
QY      662  CAGTTTCAAGATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAAACCTAAACACTCTT 721
Db      53  IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrLeu 72
QY      722  GTTCTTCCCAAGCAGCGTGATGTGATTAACATCTCTTGTATTATCCAGCAAGGCAAGCCACC 781
Db      73  PheLeuProGlnTyrThrAspAlaAspPheIleLeuValLeuLeuSerGlyLysAlaThr 92
QY      782  GTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATCSCACTC 841
Db      93  LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112
QY      842  AGAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACCAACCAAGCCCTCAGA 901
Db      113  LysLeuProAlaGlySerIleAlaTyrPheAlaAsnArgAspAsnGlnGluProArg 132
QY      902  GTAGTAAATCTCCATGCCGCTTAAACACACCGGCGCAGTTTGAGGATTTCTCCCGGCG 961
Db      133  ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152
QY      962  AGCAGCCGAGACCAATCATCTACTTGCAGGCTTTCAGCAGAAATACCTTGAGGCGCCGCC 1021
Db      153  GlyThrGlnAsnGlnLysSerSerLeuSerGlyPheSerLysAsnIleLeuGluAlaIle 172
QY      1022  TTCATGCGGAATTCATAGATACGGAGGTGCTGTTAGAGAGAAATGCAGGAGGTGAG 1081
Db      173  PheAsnThrAsnTyrGluGluIleGluLysValLeuLeuGluGlnGln-----Glu 189
QY      1082  CAAGGAGAGAGGGCAGCGGAGTGGAGTACTCGGAGTAGTGAG---AACAAATCAAGGA 1138
Db      190  GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGlnIleAsnGluGluAsn 209
QY      1139  GTGATAGTCAAAAGTCTCAAAGGAGCAGCTTCAAGAACTTACTTAAGCAGCTAAATCCGCTC 1198
Db      210  ValIleValLysValSerArgAspGlnIleGluLeuSerLysAsnAlaLysSerSer 229
QY      1199  TCAAGAAAGGCTCCGAAAGAGAGGAGATATACCAACCAACTCAACTTGAGAGAAAGGC 1258
Db      230  SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246
QY      1259  GAGCCCGATCTTTCTAACAACTTTGGGAGAGTTATTGTAGGTGAAGCCAGACAGAGAAC 1318
Db      247  AsnProIleTyrSerAsnLysPheGlyLysPheGluIleThrProGlu---LysAsn 265
QY      1319  CCCAGCTTCAGGACCTGGACATGTGCTACCTGCTGTGTAGAGATCAAAAGAGAGGACTTTG 1378
Db      266  GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysValGlySerLeu 285
QY      1379  ATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCGTCGTCACAAAGGAAGACTGGA 1438
```

Db 286 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 305
QY 1439 AACCTTTAACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGCGCGGGAAGAGAG 1498
Db 306 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 321
QY 1499 GAGGACGAAGACCAAGAGAGGAGGAAGTAACAGAGAGTGCGTAGGTACACAGCGAGG 1558
Db 322 AsPlyeGluGluGlnGluGluGluGluThrSerLysGlnValGlnLeuTyrArgAlaLys 341
QY 1559 TTGAAGGAAGCGCATGTTTCATCATGCGACGAGCTCATCCAGTAGCCATCAACGCTTCC 1618
Db 342 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 361
QY 1619 TCCGAATCCATCTGCTTGGCTGATCAACGCTGAAACCAACACAGAAATCTTCCTT 1678
Db 362 SerAspLeuAsnLeuIleGlyLeuGlyIleAsnAlaGluAsnGluArgAsnPheLeu 381
QY 1679 GCAGGTGATAAGACATGTGATAGACCATAGACAGCAAGCAAGCAAGATTAGCATTC 1738
Db 382 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 401
QY 1739 CTGGGTGCGGTCAACAAGTTGAGAAG 1765
Db 402 ProGlySerSerHisGluValAspArg 410

RESULT 6
US-07-955-905A-2
; Sequence 2, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-955-905A-2

Alignment Scores:
Pred. No.: 1,75e-69 Length: 566
Score: 811.00 Matches: 195
Percent Similarity: 50.95% Conservative: 127
Best Local Similarity: 30.85% Mismatches: 218
Query Match: 22.77% Indels: 92
DB: 1 Gaps: 19

US-10-728-323-1 (1-2032) x US-07-955-905A-2 (1-566)
QY 65 TCTCCACTGATGCTGTTGTAGGATCTTGTCTGGCTTCAATGCTTCAACGGCATGCC 124
Db 6 SerProPheIleValIleLeuPheSerLeuLeuLeuSerPheAlaLeuLeuCysSerGly 25
QY 125 AGTCAATCACTTACCAAGAAACACAGAACCCCTGCGCCAG-----AGGTGC 175
Db 26 ValSerAlaTyrGlyArgLysGlnTyrGluArgAspProArgGlnGlnTyrGluGlnCys 45
QY 176 CTCGAGATGTCAACAGAA-----CCGGATGACTTGAACAAAGCATGCCAGTCTCG 232
Db 176 CTCGAGATGTCAACAGAA-----CCGGATGACTTGAACAAAGCATGCCAGTCTCG 232

Db 46 GlnArgArgCysGluSerGluAlaThrGluGluArgGluGlnGlnCysGluGlnArg 65
QY 233 TGCACCAAGCTCGAGTATGATCTCTGTGTGTCTATGATCTCTCGAGGACACACTGCGACC 292
Db 66 CysGluArg---GluTyrLysGluGln----- 73
QY 293 ACCAACCAAGCTTCCCTCCAGGGAGCGGACACGCTGGCGCGCAACCCCGGAGACTACGAT 352
Db 74 -----GlnArgGlnGlnGlnGluGluGluGlnArgGlnTyrGln----- 86
QY 353 GATGACCGCGCTCAACCCCGAAGAGAGGAGCGCGCTGGGACCGACTGGACCGAGG 412
Db 87 -----GlnCysGlnGlnArgCysGlnGlnGlnGln-----GlyGlnArg 100
QY 413 GAGCTGAA-----AGAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGG----- 460
Db 101 GlnGlnGlnCysGlnArgLysCysTyrGluGlnTyrLysGluGlnGluArgGlyGlu 120
QY 461 CGACCAAGTCAATCAGCAGCCCGGAAATAGCCCGGAGGAGGAGGAGAGAGAGAGAGAG 520
Db 121 HisGluAsnTyrHisAsnHisLysLysAsnArgSerGluGluGluGlnGlnGln--- 139
QY 521 TGGGGAACACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTTACTTTC 580
Db 140 -----ArgAsnAsnProTyrTyrPhe 146
QY 581 CCGTCAAGCGG---TTTAGACCCGCTACGGGAACCAACCGGTAGGATCCGGTCTCTG 637
Db 147 ProLysArgArgSerPheGlnThrArgPheArgAspGluGluGluAsnPheLysIleLeu 166
QY 638 CAGAGTTTGACCAAGGTCAAGGAGTTTCAGATCTCCAGAAATCCAGAAATCCAGTATTGTG 697
Db 167 GlnArgPheAlaGluAsnSerProLeuLysGlyIleAsnAspTyrArgLeuAlaMet 186
QY 698 ATCGAGGCGCAACCTAAACACTCTGTCTTCTCCCAAGCAGCTGCTGCTGATAACACTCTT 757
Db 187 PheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAlaIleTyr 206
QY 758 GTTATCCAGCAGGCGCAACCGTACCGTAGCAAAATGGCAATTAACAAGAAGAGCTTT 817
Db 207 PheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGluSerTyr 226
QY 818 ATCTTGTAGCGGGCCATGCTCAGATCCCATCCGTTTCATTTCTTCTACATCTTGAAC 877
Db 227 AsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrValValSer 246
QY 878 CGCCATGACAACAGAACTCAGAGTAGTAAATCTCCATCCCGTAAACACACCCCGC 937
Db 247 GlnAspAsnGlnGlyLysLeuThrIleAlaValLeuAlaLeuProValAsnSerProGly 266
QY 938 CAGTTTGAGGATTTCTCCCGCGAGCAGCGAGACCAATCATCTTCTTTCAGGGCTTC 997
Db 267 LysTyrGluLeuPhePheProAlaGlyAsnAsnLysProGluSerTyrTyrGlyAlaPhe 286
QY 998 AGCAGGAATACGTTGGAGCGCCTTCAATGCGGAATTCATGATGATGAGAGGGGTGCTG 1057
Db 287 SerTyrGluValLeuLeuThrValPheAsnThrGlnArgGluGlyLysLeuGluLeuLeu 306
QY 1058 TTAGAAGAGAATGTCAGGAGGTGAGCAAGAGGAGAGGCGAGCGAGTGGAGTACTCGG 1117
Db 307 -----GluGlnArgGlyGlnLysArg----- 314
QY 1118 AGTAGTGAGAACATGAAGAGTGTAGTCAAGGTGTCAAAGGAGCAGCTTGAAGAACTT 1177
Db 315 ---GlnGlnGlnGlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIle 333
QY 1178 ACTAAGACGCTAAATCCGTCTCAAGAAAGGCTCCGAGAGAGAGGAGAGATATACCAAC 1237
Db 334 SerGlnGlnAlaThrSerProArgHisArgGlyGlyGluArgLeuAla----- 349
QY 1238 CCAATCACTTGAGAGAGGCGCGCATCTTCTTAACAACCTTTGGGAGTATTGTAG 1297
Db 350 ---IleAsnLeuLeuSerGlnSerProValTyrSerAsnGlnAsnGlyArgPhePheGlu 368

```
QY 1298 GTGAAGCCACACAGAAGAACCCACAGCTTCAGGACCTGGACATGATGCTCACCTGTGTA 1357
Db 369 AlaCysProGluAspPheSer---GlnPheGlnAsnMetAspValAlaValSerAlaPhe 387
QY 1358 GAGATCAAAAGAGGAGCTTTGATGTCCACACTTCAACTCAAGGCCATGTTATCGTC 1417
Db 388 LysLeuAsnGlnGlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValVal 407
QY 1418 GTCGTCAACAAGGAACTGGAAACCTTGAACTCGTGGCTGTAAGAAAAGAGCAACA--- 1474
Db 408 PheValThrAspGlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSer 427
QY 1475 -----CAGAGGGGACGGCGGGAAGAGGAGGACGACGACGACCAAGAGAGGAGGA 1525
Db 428 GlnGlySerGlnSerGlyArgGlnAspArgGluGlnGluGluSerGluGluGlu 447
QY 1526 AGTAACAGAGAGGTGGTGTAGGTACACAGCGAGGTGGAAGAGGCGATGTTCATCATG 1585
Db 448 ThrPheGlyGluPheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAla 467
QY 1586 CCAGGAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA-----CTCCATCTGCTTGC 1639
Db 468 ProAlaGlyHisAlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAla 487
QY 1640 TTCGGTATCAACGGCTGAAACCAACACAGAAATCTTCTTTCAGGTGATAGGACAAATGTG 1699
Db 488 PheGlyLeuAsnAlaGlnAsnAsnGlnArgIlePheLeuAlaGly---LysLysAsnLeu 506
QY 1700 ATAGACCAGATAGACAGACGAGGAGGATTTAGCATTCCTCGGTGCGGTGAAACAAGTT 1759
Db 507 ValArgGlnMetAspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuVal 526
QY 1760 GAGAAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAATCTCAA 1819
Db 527 AspAsnIlePheAsnAsnProAspGluSerTyrPheMetSerPheSerGlnArgGln 546
QY 1820 TCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGAA 1879
Db 547 -----ArgArgAspGlu 550
QY 1880 AACCAAGGAGGAGGTCCACTCTCTTTCAATTTTG 1915
Db 551 ArgArgGly-----AsnProLeuAlaSerIleLeu 560
RESULT 7
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955, 905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Theobroma cacao
; FEATURE:
; NAME/KEY: Protein
```

```
; LOCATION: 1..566
; OTHER INFORMATION: /note= "67 kD Precursor Protein"
US-07-955-905A-22
Alignment Scores:
Pred. No.: 1,75e-69 Length: 566
Score: 811.00 Matches: 195
Percent Similarity: 50.95% Conservative: 127
Best Local Similarity: 30.85% Mismatches: 218
Query Match: 22.77% Indels: 92
DB: 1 Gaps: 19
US-10-728-323-1 (1-2032) x US-07-955-905A-22 (1-566)
QY 65 TCTCCACTGATGCTCTGCTAGGGATCTGTCTCGGCTTCAGTTTCTGCAACGCATGCC 124
Db 6 SerProPheIleValIlePheSerLeuLeuLeuSerPheAlaLeuLeuCysSerGly 25
QY 125 AAGTCATCACCTTACCAGAAGAAAACAGAGAACCCCTGCGCCACG-----AGGTGC 175
Db 26 ValSerAlaTyrGlyArgLysGlnTyrGluArgAspProArgGlnGlnTyrGluGlnCys 45
QY 176 CTCGAGTGTGTCACAGGAA---CCGATGACTTGAAGCAAAAGCGATGCGAGTCTCGC 232
Db 46 GlnArgArgCysGluSerGluAlaThrGluArgGluGlnGluGlnCysGluGlnArg 65
QY 233 TGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACC 292
Db 66 CysGluArg---GluTyrLysGluGln-----73
QY 293 ACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTTGCGCCGCCAACCCGAGACTACGAT 352
Db 74 -----GlnArgGlnGlnGluGluLeuGlnArgGlnTyrGln-----86
QY 353 GATGACCCCGCTCAACCCCGAAGAGAGAGAGAGAGGCCCATGGGACACAGCTGGACCGAGG 412
Db 87 -----GlnCysGlnGlyArgCysGlnGlnGlnGln-----GlyGlnArg 100
QY 413 GAGCGTGAA-----AGAGAAGAGACTGGAGACAACCAAGAGAAGATTGGAGG-----460
Db 101 GluGlnGlnGlnCysGlnArgLysCysTrpGluGlnTyrLysGluGlnGluArgGlyGlu 120
QY 461 CGACCAAGTCTATCAGACGCCACCGAAAAATAAGGCCCGCAAGGAAGAGAGGAAACAAGAG 520
Db 121 HisGluAsnTyrHisAsnHisLysLysAsnArgSerGluGluGluGluGlnGln---139
QY 521 TGGGGAACACACAGGTAGCCATGTGAGGGAAGAAAACATCTCGGAACAACCTTCTACTTTC 580
Db 140 -----ArgAsnAsnProTyrTyrPhe 146
QY 581 CCGTCAAGGCGG---TTTAGCACCCGCTACGGGAACCAACCGGTAGGTACCGGTCTCTG 637
Db 147 ProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPhenylIleLeu 166
QY 638 CAGAGGTTTGACCAAGGTCAGGCGAGTTTCAGAAATCTCCAGAATCCACGATATTGTGCAG 697
Db 167 GlnArgPheAlaGluAsnSerProLeuLysGlyIleAsnAspTyrArgLeuAlaMet 186
QY 698 ATCAGAGGCAACCTAAACACTCTTGTCTTCTCCAGACGCGCTGATGCTGTATAACATCTT 757
Db 187 PheGluAlaAsnProAsnThrPheIleLeuProHisHisCysAspAlaIleTyr 206
QY 758 GTTATCCAGCAAGGCAAGCCAGTACCGTACCGTAGCMAATGCAATAACAGAAAGAGCTTT 817
Db 207 PheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGluSerTyr 226
QY 818 AATCTTGACGAGGCGCATGCTCAGAAATCCCATCCCGTTCCTTCTTACATCTTGAC 877
Db 227 AsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrValValSer 246
QY 878 CGCCATGACACACAGAACCTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCCGCG 937
Db 247 GlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSerProGly 266
```



```
QY 1067 AATGCAAGGAGTGTGACCAAGAGGAGAGGAGGCGAGGCGATGGAGTACTCGGAGTAGTGAG 1126
Db |||:||||
209 Glu-----GlyGlnGln----- 212
QY 1127 AACATGAAGGAGTGTATGATCAAGTGTCAAGGAGCACGTTGAAGAACTTACTAAGCAC 1186
Db |||:|||||
213 -----GluGlyValIleValAenIleAsePserGluGlnIleValLeuSerIysHis 230
QY 1187 GCTAATCGTCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCCAATCAAC 1246
Db |||:|||||
231 AlaLysSerSerArgLys-----Ser 238
QY 1247 TTGAGAGAAGGCGGAGCCGCTTTCTTAACAACCTTTGGGAAGTTATTTGAGGTGAAGCCA 1306
Db |||:|||||
239 LeuSerLysGlnAsePasnThrIleGlyAenGluPheGlyAsePasnLeuThrGlu---ArgThr 257
QY 1307 GACAGAAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGATCAAA 1366
Db |||:|||||
258 Asp-----AsePasnLeuAsePasnValIleSerIleGluMetGlu 271
QY 1367 GAAGAGCTTGTGCTCCACACTCAACTCAAGGCCATGTTATCTGCTGCTCAAC 1426
Db |||:|||||
272 GluGlyAlaLeuPheValProHisTyrTyrSerLysAlaIleValIleLeuValAsePasn 291
QY 1427 AAAGGAACCTGGAACCTTCAACTCGTGGCTGTAAGAAAAGAGCAACACAGAGGGGAGCG 1486
Db |||:|||||
292 GluGlyGluAlaHisValGluLeuValGly-----ProLysGlyAsePasn 305
QY 1487 CGGAAGAAGAGGAGGAGCAAGACGACGAGAGGAGGAGGAAGTAACAGAGAGTGGTAGG 1546
Db |||:|||||
306 LysGluThrLeuGluTyrGluSer----- 313
QY 1547 TACACAGCAGGTTGAAGGAGCGGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCC 1606
Db |||:|||||
314 TyrArgAlaGluLeuSerLysAsePasnValPheValIleProAlaIleTyrProValAla 333
QY 1607 ATCAAGCGTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAAGCGTGAAACCAACAC 1666
Db |||:|||||
334 IleLysAlaThrSerAsePasnValAsePasnPheThrGlyPheGlyIleAsePasnAsePasn 353
QY 1667 AGAATCTTCTTGGAGGTGATGAAGCAATGTATGATACCAAGATAGAGAG-----CAA 1720
Db |||:|||||
354 ArgAsePasnLeuAlaGlyLysThrAsePasnValIleSerIleGlyArgAlaLeuAsePasn 373
QY 1721 GCGAAGGAT-----TTAGCATTCCTGGTGGTGGTGAACAAGTTGAGAGCTCATC 1771
Db |||:|||||
374 GlyLysAsePasnValLeuGlyLeuThrPheSerGlySerGlyAsePasnValMetLysLeuIle 393
QY 1772 AAAAAACAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCCG 1831
Db |||:|||||
394 AsePasnLysGlnSerGlySerTyrPheValAsePasnAlaHis----- 405
QY 1832 TCGTCTCTGAGAAGAGTCTCTGAGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGG 1891
Db |||:|||||
406 -----HisHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 415
QY 1892 AAGGTT 1897
Db |||:|||||
416 LysGly 417
```

RESULT 9

```
US-07-955-905A-23
; Sequence 23, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..587
; OTHER INFORMATION: /note= "Vicilin from G. hirsutum"
US-07-955-905A-23
```

Alignment Scores:

Pred. No.:	3,53e-66	Length:	587
Score:	777.00	Matches:	190
Percent Similarity:	51.36%	Conservative:	112
Best Local Similarity:	32.31%	Mismatches:	208
Query Match:	21.81%	Indels:	78
DB:	1	Gaps:	15

US-10-728-323-1 (1-2032) x US-07-955-905A-23 (1-587)

```
QY 104 TCAGTTCTTCCAAACGATGCCAAGTCATCACCTTACCAGAAAGAAACAGAGAACCCCTGC 163
Db |||:|||||
64 SerCysLysSerGlnTyrGlyGluLysAsePasnGlnArgHisArgProGluAsePasnPro-- 82
QY 164 GCCCAGAGTCCCTCCAGAGTGTTCACAGGAA-----CCGATGACTTGAAGCAA 214
Db |||:|||||
83 ---GlnArgArgTyrGlyGluGluCysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 101
QY 215 AAGGCATCCGAGTCTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCT 274
Db |||:|||||
102 ProGlnCysGlnGlnArgCysIleLys----- 110
QY 275 CGAGGACACTGGCACCACCAACCAACGTTTCCCTCCAGGGGAGCGGACAGTGGCCGC 334
Db |||:|||||
111 -----ArgPheGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 121
QY 335 CAACCCGAGACTAGATGATGAC---CGCGTCACACCCCGAAGAGAGAGAGAGAGAGAGAG 391
Db |||:|||||
122 GlnPheGlnGluCysGlnGlnHisCysHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 141
QY 392 TGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451
Db |||:|||||
142 GlnCysValAlaGluCysArgGluArgTyrGlnGluAsePasnProTyrArgGluArgGlu 161
QY 452 GATTGAGGCGACCAAGTCATCAGCAGCCACGGAATAAGGCCCGGAGAGAGAGAGAGAG 511
Db |||:|||||
162 Glu-----GluAlaGluGlu 167
QY 512 GAACAGAGTGGGNAACACACAGGTAGCATGTGAGGGAAGAAACATCTCGGACACACCT 571
Db |||:|||||
168 GluThrGluGluGly-----GluGlnGlnGlnSerHisAsePasnPro 180
QY 572 TTCTACTTCCCGTCAAGCGGTTTAGCACCGCTACGGGAACCAACAAACCGTAGGATCCGG 631
Db |||:|||||
181 PheHisPheHisArgArgSerPheGlnSerArgPheArgGluGluHisGlyAsePasnPhe 200
QY 632 GTCCTGCAAGGTTTGACCAAGAGTCAAGGCGAGTTTTCAGATCTCCAGAAATCACCGTATT 691
Db |||:|||||
201 ValLeuGlnArgPheAlaSerArgHisProIleLeuArgGlyIleAsePasnGluPheArgLeu 220
QY 692 GTGCAGATCGAGGCGCAACCTTAACACTCTTCTTCCCAAGAGAGAGAGAGAGAGAGAG 751
Db |||:|||||
221 SerIleLeuGluAlaAsePasnProAsePasnThrGluValLeuProHisCysAsePasnGlu 240
```



```
Db 110 ArgValGluThrGluGlyGlySerIleArgValLeuLysPheThrGluLysSerLys 129
QY 662 CAGTTTCCAGAACTCCAGAAATCACCCTATGTGCAGATCAGGCCAAACCTAACACTCTT 721
Db 130 LeuLeuGlnGlyLeuGluAsnPheArgLeuAlaIleLeuGluAlaArgAlaHisThrPhe 149
QY 722 GTTCTTCCCAAGCAGCTGATGCTGATAACATCTCTTTTATCCAGCAAGGCGCAAGCCACC 781
Db 150 ValSerProArgHisPheAspSerGluValValPheAsnIleLysGlyArgAlaVal 169
QY 782 GTGACCGTAGCAAAATGCCAATAACAGAAAGAGCTTTAATCTTTCAGCAGGCGCCATGCATC 841
Db 170 LeuGlyLeuValArgLeuSerGluThrGluLysIleThrLeuGluProGlyAspMetIle 189
QY 842 AGAATCCCATCCGTTTTCATCTTCTACATCTTGAACCGCCATGCACACCAACCTCAGA 901
Db 190 HisIleProAlaGlyThrProLeuTyfIleValAsnArgAspGluAsnGluLysLeuLeu 209
QY 902 GTAGCTAAATCTCCATGCCGTTAACACACCCGGCCAGTTTGAGGATTTCTCCCGGGC 961
Db 210 LeuAlaMetLeuHisIleProValSerThrProGlyLysPheGluGluPheGlyPro 229
QY 962 AGCAGCCGAGACCAATCTCTACTTTCAGGCTTCAGCAGGAATACGTTGAGGCCGCC 1021
Db 230 GlyGlyArgAspProGluSerValLeuSerAlaPheSerTyrAsnValLeuGlnAla 249
QY 1022 TTCATGCGGAATTCATGATGATCAGGAGGCTGCTGTAGAGAGAAATCCAGGAGGTGAG 1081
Db 250 LeuGlnThrProLysGlyLysLeuGluArgLeuPhe 261
QY 1082 CAAGAGGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAATGAAGGAGTG 1141
Db 262 262
QY 1142 ATAGTCAAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCA 1201
Db 269 IlePheLysIleSerArgGluArgValArgAlaLeu 283
QY 1202 AGAAAGGCTCC-----GAAGNAGGGGAGATATCACCAACCCATCAACTTGTAGAGNA 1255
Db 284 LysLysSerSerTyrTrpProPheGlyGlyLysLeuSerLysAlaGlnPheAsnIlePheSer 303
QY 1256 GCGAGCCCGATCTTTCTAACTTTGGAAGTATTATTTGAGTGAAGCAGACAGAAAG 1315
Db 304 LysArgProThrPheSerAsnGlyTyrGlyArgLeuThrGluValGlyProAspAspGlu 323
QY 1316 AACCCCGACTTCAGACCTCGACATGATGCTCACCCTGTGTAGAGATCAAGAAAGAGCT 1375
Db 324 LysSerTyrLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSer 343
QY 1376 TTGATGCTCCCACTTCAACTCAAGGCCATGGTTATCGTGTCTCAACAAAGAACT 1435
Db 344 MetSerThrIleHisTyrAsnSerHisAlaThrLysIleAlaLeuValMetAspGlyArg 363
QY 1436 GGAACCTTGAACCTCGTGTCTGAAGAAAGAGCAACACAGAGGGGCGCGGGAAGAA 1495
Db 364 GlyHisLeuGlnIleSerCysProHisMetSerSerArg 376
QY 1496 GAGGAGGACGAAGCAAGAGGAGGAGGAAGTAACAGAGAGTGCGTAGGTACACAGCG 1555
Db 377 -----SerAspSerLysHisAspLysSerSerProSerTyrHisArgIleSerAla 393
QY 1556 AGTTTGAAGGAGGCGATGTTTCATCATCGCAGCAGCTCATCCAGTACCCATCAACGCT 1615
Db 394 AspLeuLysProGlyMetValPheValProGlyHisProPheValThrIleAla 413
QY 1616 TCTCCGAACTCCATCTGCTT-----GGCTTCGGTATCAACGCTGAAACCAACACAGAGA 1669
Db 414 SerAsnLysGluAsnLeuLeuIleCysPheGluValAsnValArgAspAsnLysLys 433
QY 1670 ATCTTCTTGGAGGTGATTAAGACAAATGTGATAGCACCATAGAGAGCAAGCGGAGGAT 1729
```

```
Db 434 PheThrPheAlaGly---LysAspAsnIleValSerSerLeuAspAsnValAlaLysGlu 452
QY 1730 TTAGCATTTCCCTGGGTCCGGTGCAACAACCTTGAGAAGCTCATCAAAACCCAGAAGAACTCT 1789
Db 453 LeuAlaPheAsnTyrProSerGluMetValAsnGly----ValSerGluArgLysGluSer 471
QY 1790 CACTTT 1795
Db 472 LeuPhe 473

RESULT 11
US-09-424-283-1
; Sequence 1, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-1

Alignment Scores:
Pred. No.: 4,738-46 Length: 524
Score: 569.00 Matches: 163
Percent Similarity: 44.03% Conservative: 106
Best local Similarity: 26.68% Mismatches: 208
Query Match: 15.97% Indels: 134
DB: 4 Gaps: 17

US-10-728-323-1 (1-2032) x US-09-424-283-1 (1-524)
QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGTAGGGATCCTTGTCTGCTTTCAGTT 109
Db 3 MetArgThrLysLeuSerLeuAlaIlePhePhePheLeuLeuAlaLeuPheSer--- 21
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAGAAACAGAGAACCCCTGCCCCAG 169
Db 22 AsnLeuAlaPheGlyLysCysLysGluThrGluValGluGluAspProGluLeuVal 41
QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGAT-----GACTTGAAGCAAAAGCATGC 223
Db 42 ThrCysLysHisGlnCysGlnGlnGlnGlnGlnTyrThrGluGlyAspLysArgValCys 61
QY 224 GAGTCTCCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACAC 283
Db 62 LeuGlnSerCys----- 65
QY 284 ACTGCGACCAACAACCAACGTTCCCTCCAGGGAGCGGACACGTTGGCGGCCAACCCCGA 343
Db 65 ----- 65
QY 344 GACTACGATGATACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGGACCGT 403
Db 66 -----AspArgTyrHisArgMetLysGlnGlu----- 74
QY 404 GGACCGAGGGAGCGGTGAAGAAGAAAGACTGGAGACAACCAAGAGAAGATTGGAGCGGA 463
Db 75 -----ArgGluLysGlnIleGlnGluThrArgGluLysLysGluGlu----- 89
QY 464 CCAAGTCATCAGCACCCACGAAATAAGGCCCGGAGGAGGAGGAGGAGGAGGAGTGG 523
Db 90 -----GluSerArgGluArgGluGlu----- 97
```

QY 524 GGAACACCGTAGCCATGTGAGGGAAGAACATCTCGGAACACCCCTTTCTACTTC--- 580
Db : : : : :
QY 98 -----GlnGlnGlnHisGlnGlnAspGluAsnProTyrIlePheGlu 113
Db : : : : :
QY 581 CCGTCAAGGGGGTTTAGCACCCGCTACCGGAACCAAAACGGTAGGATCCGGGCTCTGCAG 640
Db : : : : :
QY 114 GluAspLysAspPheGluThrArgValGluThrGluGlyArgIleArgValLeuLys 133
Db : : : : :
QY 641 AGTTTGACCAAAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCCACCGTATGTGCGAGATC 700
Db : : : : :
QY 134 LysPheThrGluLysSerLysLeuLeuGlnGlyIleGluAsnPheArgLeuAlaIleLeu 153
Db : : : : :
QY 701 GAGGCCAAACCTAACACTTGTCTTCCNAGCAGCTGATGCTGATAACATCTTGT 760
Db : : : : :
QY 154 GluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluValValPhePhe 173
Db : : : : :
QY 761 ATCCAGCAAGGGCAAGCCCGTACCGTAGCAAAATGGCAATACACAGAGAGCTTTAAT 820
Db : : : : :
QY 174 AsnIleLysGlyArgAlaValLeuGlyLeuValSerGluThrGluLysIleThr 193
Db : : : : :
QY 821 CTGACGAGGGCCATGACCTCAGAAATCCCATCCGTTTTCATTTCTACATCTTGAACCGC 880
Db : : : : :
QY 194 LeuGluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArg 213
Db : : : : :
QY 881 CATGACACACGACCTCAGAGTAGCTAAATCTCCATGCC-----GTTACACACCC 934
Db : : : : :
QY 214 AspGluAsnAspLysLeuPheLeuAlaMetLeuHisIleProValSerValSerThrPro 233
Db : : : : :
QY 935 GGCCAGTTTGGAGATTTCTTCCCGGCGAGCGACCAATCATCTCTCTTGGAGGC 994
Db : : : : :
QY 234 GlyLysPheGluPhePheAlaProGlyArgAspProGluSerValLeuSerAla 253
Db : : : : :
QY 995 TTCAGCAGGAATAGTTGGAGCGCTTCAATCGGAAATTCATGAGATACGAGGGTG 1054
Db : : : : :
QY 254 PheSerTrpAsnValLeuGlnAlaLeuGlnThrProLysGlyLysLeuGluAsnVal 273
Db : : : : :
QY 1055 CTGTTAGACAGATGACGAGGTGACCAAGGAGAGAGCGGCGAGCGGATCGTACT 1114
Db : : : : :
QY 274 Phe----- 274
QY 1115 CGGAGTAGTGAGAACATGAAGAGTGATAGTCAAAGTGTCAAAGGAGACGTTGAAGAA 1174
Db : : : : :
QY 275 -----AspGlnGlnAsnGluGlySerIlePheArgIleSerArgGluGlnValArgAla 292
Db : : : : :
QY 1175 CTACTAAGCACGCTAAATCCGCTCAAGAAAGGCTCC-----GAAGAAGAGGAGAT 1228
Db : : : : :
QY 293 Leu-----AlaProThrLysLysSerSerTrpTrpProPheGlyGlyGlu 307
Db : : : : :
QY 1229 ATCACCAACCAATCAACTTGAAGAGGCGAGCGCGATCTTCTTAACAATTTGGGAAG 1288
Db : : : : :
QY 308 SerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGlyTyrGlyArg 327
Db : : : : :
QY 1289 TTATTTGAGTGAAGCCA---GACAAGAAGAACCCCGCTCAGGACCTGGACATGATG 1345
Db : : : : :
QY 328 LeuThrGluValGlyProAspAspArgGlyLysSerTrpLeuGlnArgLeuAsnLeuMet 347
Db : : : : :
QY 1346 CTCACCTGTGTAGAGATCAAGAAGAGGAGCTTGTATGCTGCCACATTTCAACTCAAAAGGCC 1405
Db : : : : :
QY 348 LeuThrPheThrAsnIleThrGlnArgSerMetSerThrIleHisTyrAsnSerHisAla 367
Db : : : : :
QY 1406 ATGGTTATCTGCTGCTCAACAAGGAACTGGAAACCTTGAACCTGTGGCTGTGAAGAAA 1465
Db : : : : :
QY 368 ThrLysIleAlaLeuValIleAspGlyArgGlyHisLeuGlnIleSerCysProHisMet 387
Db : : : : :
QY 1466 GAGCAACAACAGAGGGGACCGCGGAAGAAGAGGAGGAGCAAGACGAAGAAGAGGAGGA 1525
Db : : : : :
QY 388 SerSerArgSerSerHisSerLysHisAspLys----- 398
Db : : : : :
QY 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGTGAAGAGGCGATGTGTTCATCATG 1585
Db : : : : :
QY 399 SerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPheValVal 418
Db : : : : :

QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACCTCCATCTGCTT-----GGC 1639
Db : : : : :
QY 419 ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuMetIleCys 438
Db : : : : :
QY 1640 TTCGGTATCAACGCTCAAAACACACACAGAAATCTTCTTCAGGTGATAGGACAATGTG 1699
Db : : : : :
QY 439 PheGluValAsnAlaArgAspAsnLysPheThrPheAlaGly---LysAspAsnIle 457
Db : : : : :
QY 1700 ATAGACAGATGAGGAAGCAAGCAAGGATTTAGCATTCCTCGGTGCGGTGAACAAGTT 1759
Db : : : : :
QY 458 ValSerSerLeuAspAsnValAlaLysGluLeuAlaPheAsnTyrProSerGluMetVal 477
Db : : : : :
QY 1760 -----GAGAAGCTCATC----- 1771
Db : : : : :
QY 478 AsnGlyValPheLeuLeuGlnArgPheLeuGluArgLysLeuIleGlyArgLeuTyrHis 497
Db : : : : :
QY 1772 -----AAAAACCAAGGAATCTCACTTT 1795
Db : : : : :
QY 498 LeuProHisLysAspArgLysGluSerPhePhe 508
Db : : : : :
RESULT 12
US-09-323-195A-18
; Sequence 18, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Perrera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Picea glauca
US-09-323-195A-18
Alignment Scores:
Pred. No.: 3,57e-45 Length: 448
Score: 559.50 Matches: 134
Percent Similarity: 50.12% Conservative: 82
Best Local Similarity: 31.09% Mismatches: 178
Query Match: 15.71% Indels: 37
DB: 4 Gaps: 10
US-10-728-323-1 (1-2032) x US-09-323-195A-18 (1-448)
QY 545 AGGGAAGAAACATCTCGGAACCAACCCCTTTCTACTTCCCTCAAGCGGTTTAGCACCCGC 604
Db : : : : :
QY 44 ArgGluGluArgGluGluAsnProTyrValPheHisSerAspSerPheArgThrArg 63
Db : : : : :
QY 605 TACGGGAACCAAAACGGTAGGATCCGGGTCCTCGAGAGTTTGACCAAGGTCACAGGCAG 664
Db : : : : :
QY 64 AlaserSerGluAlaGlyGluIleArgAlaLeuProAsnPheGlyGluValSerGluLeu 83
Db : : : : :
QY 665 TTTCAGATCTCCAGATCACCAGTATTGTGCAGATCGAGGCCCAACCTTAACACTCTTGT 724
Db : : : : :
QY 84 LeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThrValMet 103
Db : : : : :
QY 725 CTTCCCAAGCAGCGCTGATCTGATAACATCTTGTATTCCAGCAGGGCAAGCCACCGTG 784
Db : : : : :
QY 104 LeuProHisTyrIleAspAlaThrIleLeuTyrValThrArgGlyArgGlyTyrIle 123
Db : : : : :
QY 785 ACCGTAGCAATGGCAATTAACAAAGAGCTTTAATCTTGACGAGGGCCATGCACCTCAGA 844
Db : : : : :
QY 124 AlaTyrValHisGlnAsnGluLeuValLysArgLysLeuGluGluGlyAspValPheGly 143
Db : : : : :
QY 845 ATCCCATCCGGTTTCATTTCTTACATCTTGAACCCCATGACCAACAGAACCTCAGAGTA 904
Db : : : : :

```
Db 144 ValProSerGlyHisThrPheTyrLeuValAsnAsnAspHisAsnThrLeuArgIle 163
QY 905 GCTAAATCTCCATGCGCGTTAACACA--CCGGCCAGTTTGAGGATTTCTCCGGCG 961
Db 164 AlaSerLeuValArgProValSerThrValArgGlyGluTyrGlnProPheTyrValAla 183
QY 962 AGCAGCCGAGACCAATCATCTTCTGACGGCTTCAGCAGGAATACGTTGGAGCGCGC 1021
Db 184 GlyGlyArgAsnProGlnThrValTyrSerAlaPheSerAspValLeuGluAlaAla 203
QY 1022 TTCATTCGGAATTCATGAGTACGAGGCTGCTGTAGAGAGAAATCCAGGAGGTGAG 1081
Db 204 PheAsnThrAsnValGlnGlnLeuGluArgIlePhe-----GlyGly---- 217
QY 1082 CAAGAGGAGAGGGCAGAGCGGTGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTG 1141
Db 218 -----HisLysSerGlyVal 222
QY 1142 ATAGTCAAAAGTGTCAAAGGAGCAGTTGAAGAACTTTACTAAGCACGCTAAATCCGTCTCA 1201
Db 223 IleIleHisAlaAsnGluGlnIleArgGluMetMetArg-----LysArgGlyPhe 240
QY 1202 AGAAAGGCTCCGAAGAGAGGAGATATCACCAACCCCAATCACTTGAGAGAGGCGGAG 1261
Db 241 SerAlaGlySerMetSerAlaProGluHisProLysProPheAsnLeuArgAsnGlnLys 260
QY 1262 CCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAGCCAGCACAGAAAGACCCC 1321
Db 261 ProAspPheGluAsnGluAsnGlyArg---PheThrIleAlaGlyProLysAsnTyrPro 279
QY 1322 CAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATG 1381
Db 280 PheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeuAsnProGlySerMetThr 299
QY 1382 CTCCHCACTTCAACTCAAGGCCATGTTATCGTCGTCGTCAACAAAGAACTGGAAC 1441
Db 300 AlaProSerLeuAsnSerLysSerThrSerIleGlyIleValThrAsnGlyGluGlyArg 319
QY 1442 CTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGGAG 1501
Db 320 IleGluMetAlaCysProHisLeuGlyGlnHisGlyTyrPheSerProArgGluArgGly 339
QY 1502 GACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACACGCGAGTTG 1561
Db 340 AspGlnAsp-----IleThrTyrGlnArgValTrpAlaLysLeu 352
QY 1562 AAGGAAGGGAGTGTTCATCATGCCAGCAGCTCATCCAGTAGCCCATCAACGCTTCC--- 1618
Db 353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr 372
QY 1619 ---TCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACACACAGAATCTTC 1675
Db 373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe 392
QY 1676 CTTTCAGGTGATAAGGACAATGTGATGACCAAGATAGAGCAAGCAGGAGGATTTAGCA 1735
Db 393 LeuAlaGly---LysAsnAsnValLeuAsnThrLeuGluArgGluIleArgGlnLeuSer 411
QY 1736 TTC---CCTGGTGGGTGAACAAGTTGAGAGCTCATCAAAACACAGAAAGAACTTCAC 1792
Db 412 PheAsnValProArgGlyGluGluIleGluGluValLeuGlnAlaGlnLysAsp---Gln 430
QY 1793 TTGTGAGTGCTCGTCTCAATCTCAATCTCAA 1825
Db 431 ValIleLeuArgGlyProGlnArgArgSerArg 441
```

RESULT 13

US-09-424-283-4

; Sequence 4, Application US/09424283

; Patent No. 6437219

; GENERAL INFORMATION:

; APPLICANT: Grimes, et al.

; TITLE OF INVENTION: Sucrose binding proteins

```
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-4
Alignment Scores: 7,75e-41 Length: 409
Pred. No.: 514.50 Matches: 133
Score: 514.50 Conservative: 87
Percent Similarity: 42.72% Mismatches: 168
Best Local Similarity: 25.83% Indels: 127
Query Match: 14.44% Gaps: 10
DB: 4
```

US-10-728-323-1 (1-2032) x US-09-424-283-4 (1-409)

```
QY 71 CTGATGCTGTTGCTAGGGATCCTTGTCTCTGCTTCTGCAACGCATGCCAAGTCA 130
Db 16 LeuLeuAlaLeuIleSerAsnLeuAlaLeuGlyLysLeuLysGluThrGluValGlu--- 34
QY 131 TCACCTTACCAGAAACACAGAACCCCTCGCGCCAGAGGTGCTCCAGAGTTGTCAA 190
Db 35 -----GluAspProGluLeuValThrCysLysHisGlnCysGln 47
QY 191 CAGGAACCGGAT-----GACTTGAAGCAAAAGGCATCGGAGTCTCGTCGACCAAGCTC 244
Db 48 GlnGlnArgGlnTyrThrGluSerAspLysArgThrCysLeuGlnGlnCysAspSerMet 67
QY 245 GAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCAAGT 304
Db 68 LysGlnGlu----- 70
QY 305 TCCCTCCAGGGGAGCGACACAGCTGGCGCCCAACCCGAGACTACGATGATGACCGCGGT 364
Db 70 ----- 70
QY 365 CAACCCCGAAGAGAGAGGAGCGCGATGGGACACAGCTGGACCGAGGGAGCGTGAAGA 424
Db 71 -----ArgLysGlnVal 75
QY 425 GAAGAAGACTCGAGACAAACAGAGAGATTGGAGCGACCAAGTCAATCAGCAGCCACGG 484
Db 76 GluGluGluThrArgGluLysGluGlu-----HisGlnGluGln--- 89
QY 485 AAATAAGCCCGAAGGAAGAGAGAGAACAGAGTGGGGAACACCAAGGTAGCCATGTG 544
Db 89 ----- 89
QY 545 AGGGAAGAAACATCTCGGAAACAACCTTTCTACTTC-----CCGTCAAGCGGTTTAGCACC 601
Db 90 HisGluGluGluGluAspGluAsnProTyrValPheGluGluAspLysAspPheSerThr 109
QY 602 CGCTACGGGAACCAACGCTAGGATCGGGTCTCTGACAGAGTTTGACCAAGGTCAAGG 661
Db 110 ArgValGluThrGluGlyGlySerIleArgValLeuLysLysPheThrGluLysSerLys 129
QY 662 CAGTTTTCAGATCTCCAGATCAACCGTATTGTGCAGATCGAGGCGCAAACTTAACACTCTT 721
Db 130 LeuLeuGlnGlyIleGluAsnPheArgLeuAlaIleLeuGluAlaArgAlaHisThrPhe 149
QY 722 GTTCTTCCAGACGCTGATGATAACATCTCTGTTATCTCCAGCAAGGCAAGCCACC 781
Db 150 ValSerProArgHisPheAspSerGluValValLeuPheAsnIleLysGlyArgAlaVal 169
```

Qy	782	GTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATCGCACTC	841
Db	170	LeuGlyLeuValArgGluSerGluThrGluLysIleThrLeuGluProGlyAspMetIle	189
Qy	842	AGAATCCATCCCGTGTTCATTTCTTACATCTTTGAACCGCATGACAAACCGAACCTCAGA	901
Db	190	HisIleProAlaGlyThrProLeuTyrlleValAsnArgaspGluAsnGluLysLeuLeu	209
Qy	902	GTAGCTAAATCTCCATGCCCGTTTAAACACCCCGCGAGTTTGAGGATTTCTTCCCGCGC	961
Db	210	LeuAlaMetLeuHisIleProValSerThrProGlyLysPheGluGluPhePheGlyPro	229
Qy	962	AGCAGCGGAGGACCAATCTTACTCTCCAGGCTTCACGAGGATATCGTTGAGGCGCGCC	1021
Db	230	GlyGlyArgaspProGluSerValLeuSerAlaPheSerTrpAenValLeuGlnAla	249
Qy	1022	TTCAATCGGAAATTCATATGAGATACGGAGGTGCTGTTAGAAAGAGAATGCAGGAGGTGAG	1081
Db	250	LeuGlnThrProLysGlyLysLeuGluArgLeuPhe	261
Qy	1082	CAAGAGGAGAGCGGACGAGCGATCTCGGAGTACTCGGAGTAGTGAGACAATGAAGGATG	1141
Db	262	-----AsnGlnGlnAsnGluGlySer	268
Qy	1142	ATAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACGCTAAATCCGTCTCA	1201
Db	269	IlePheLysIleSerArgGluArgValArgAlaLeu	283
Qy	1202	AAGAAAGCGCTCC-----GAAGAAGAGGGAGATATCACCAACCCAACTCAACTGAGAGAA	1255
Db	284	LysLysSerSerTrpTrpProPheGlyGlyGluSerLysAlaGlnPheAsnIlePheSer	303
Qy	1256	GCGAGGCGCATCTTTCTAACAACTTTGGGAAGTTATTGTGAGTGAAAGCCAGACAAGAAG	1315
Db	304	LysArgProThrPheSerAsnGlyTyrglyArgLeuThrGluValGlyProAspAspGlu	323
Qy	1316	AACCCCGAGCTTCAGGACCTGGACATCATGCTCACTCTGTGTAGAGATCAAGAAGGAGCT	1375
Db	324	LysSerTrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSer	343
Qy	1376	TTGATGCTCCACATTCACACTCAAAGCGCATGTTATCGTCGTCGTCACAAAGGAAC	1435
Db	344	MetSerThrIleHisTyAsnSerHisAlaThrLysIleAlaLeuValMetAspGlyArg	363
Qy	1436	GGAAACCTTGAACTCGTGGCTGTGAAGAAAGAGGACAAACAAGAGGGGACCGCGGGAAGAA	1495
Db	364	GlyHisLeuGlnIleSerCysProHisMetSerSerArg	376
Qy	1496	GAGGAGGACGAAGAGGAAGGAGGGAAGTAACAGAGAGGTCGTAGGTACACAGCG	1555
Db	377	-----SerAspSerLysHisAspLysSerSerProSerTyHisArgIleSerAla	393
Qy	1556	AGGTTGAAGGAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCA	1600
Db	394	AspLeuLysProGlyMetValPheValProGlyHisPro	408

RESULT 14

US-09-323-195A-17	QY	1181	AAGCACGCTAAATCCGTCCTCAAGAAAGGCTCCGAAGAAGGAGATATACCAACCCA	1240
Sequence 17, Application US/09323195A				
Patent No. 6462257				
GENERAL INFORMATION:				
APPLICANT: Pullman, Gerald	Db	241	LysArgGlyPheSerAlaGluSerMetSerAlaSerGluHis-----ProLysPro	257
APPLICANT: Caliney, John				
APPLICANT: Perrera, Ranjan	QY	1241	ATCAACTTGAGAGAAGGCGAGCGCGATCTTTCTAACCACTTTGGGAAGTTATTGTAGGTG	1300
TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND				
METHODS OF USING THE SAME	Db	258	PheAsnLeuArgAsnGlnLysProAspPheGluAsnAspAsnGlyArgPheThrArgAla	277
FILE REFERENCE: IPST0009				
CURRENT APPLICATION NUMBER: US/09/323,195A	QY	1301	AAGCCAGACAGAGAAGAACCCCGACGCTTCAGGACCTTCGACATGATGCTCACTGTGTAGAG	1360
CURRENT FILING DATE: 1999-06-01				
NUMBER OF SEQ ID NOS: 19	Db	278	GlyProAsn--GluAsnProLeuLeuAspAlaValAspValThrAlaGlyPheGlyVal	296
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO 17	QY	1361	ATCAAGAAGGAGCTTTCATGCTCCACACTTCACCTCAAGAGGCCATGGTTATCGTCGTC	1420
LENGTH: 523	Db	297	LeuAsnProGlyThrMetThrAlaProSerHisAsnThrLysAlaThrSerIleAlaIle	316

QY	1421	GTCAACAAAGGAACCTGAAACCTTGAACTCGTGCTGTGAAGAAAGAGCAACAACAGAGG	1480
Db	317	ValThrGlnGluGluGlyArgIleGluMetAlaCysProHisLeuGlyGlnHisGlyTyr	336
QY	1481	GGACGGCGGGAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTG	1540
Db	337	SerSerArgArgLysGlyAspGlnGlu-----IleAsnTyr	349
QY	1541	CGTAGGTACACAGCGAGGTGAAGGAGGGGATGTGTTTCATCATGCCAGCAGCTCATCCA	1600
Db	350	GlnArgValArgAlaArgLeuArgThrGlyThrValTyrValProAlaGlyHisPro	369
QY	1601	GTAGCCATCAACGCTCTCCGAACTCCCATCTG-----CTTGGCTTCGGTATCAACGCT	1654
Db	370	IleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTyrPheAspIleAsnThr	389
QY	1655	GAACCAACACACAGAAATCTTCCTTGAGGTGATAGGACAAATGTGATAGCAGATAGAG	1714
Db	390	SerGlyAsnGluArgGlnPheLeuAlaGly---LysTyrAsnValLeuGlnThrLeuGlu	408
QY	1715	AGCAAGCGAAGGATTTAGCATTCCT---GGTCGGGTGAACAAGTTGAGAGCTCATC	1771
Db	409	LysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluValLeu	428
QY	1772	AAAAACCAAGAAGAAATCTCACTTTGTGAGTGTGCTCGTCTCAATCTCAATCTCAATCTCCG	1831
Db	429	ArgArgGlnLysAsp---GlnValIleLeuArgGlyProGlnMetGlnArg---	445
QY	1832	TCGCTCTCTGAGNAAGAGT-----CTCTGAGAAAGAGATCAAGAGGAGGAA	1879
Db	446	AspGluPro-ArgSerSerSerIleHisAlaIleIleAlaAlaArgIleMetAspAspIly	465
QY	1880	AACCAAGGAGGAGGAGGTCCACTCTTTCAATTTGAAGCTTTTAACTGAGAAATGGAGG	1939
Db	465	sThr-----ArgVal-SerProArgSerLeuPheSerIleTyrLeuArgIleSer-	481
QY	1940	CAACTTGTATGTATCGATAATAAGATCACGCTTTTGTACT-----C	1981
Db	482	-----TyrAlaTyrMetLysProLysLysCysValArgArgAlaProPheIleL	498
QY	1982	TACTATCAAAAACCTTATCAATAATAAAGCTTTGTGCGTTGTTT	2028
Db	498	euMetAsnValTyrMetSerPheAsnLysProIleValGlyLeuPhe	513
RESULT 15			
US-09-424-283-2			
; Sequence 2, Application US/09424283			
; Patent No. 6437219			
; GENERAL INFORMATION:			
; APPLICANT: Grimes, et al.			
; TITLE OF INVENTION: Sucrose binding proteins			
; FILE REFERENCE: 4630-50206			
; CURRENT APPLICATION NUMBER: US/09/424,283			
; CURRENT FILING DATE: 1999-11-19			
; PRIOR APPLICATION NUMBER: PCT/US98/10465			
; PRIOR FILING DATE: 1998-05-21			
; PRIOR APPLICATION NUMBER: US 60/047,568			
; PRIOR FILING DATE: 1997-05-22			
; NUMBER OF SEQ ID NOS: 15			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 444			
; TYPE: PRT			
; ORGANISM: Glycine max			
US-09-424-283-2			
Alignment Scores:			
Pred. No.:	1.66e-39	Length:	444
Score:	501.00	Matches:	142
Percent Similarity:	43.85%	Conservative:	97
Best Local Similarity:	26.06%	Mismatches:	132
Query Match:	14.07%	Indels:	114

DB:	4	Gaps:	14
US-10-728-323-1 (1-2032) x US-09-424-283-2 (1-444)			
QY	50	ATGAGAGGAGGTTTCTCCACTGATGCTGTGTAGGGATCCTTGTCTCGCTTCAGTT	109
Db	3	MetArgThrLysLeuSerLeuAlaIlePhePhePheLeuLeuAlaLeuPheSer---	21
QY	110	TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAAGAAAACAGAGAAACCCCTGCGCCACG	169
Db	22	AsnLeuAlaPheGlyLysCysLysGluThrGluValGluGluAspProGluLeuVal	41
QY	170	AGGTCCCTCCAGAGTTGTCAACAGAACCGGAT-----GACTTGAACAAAAGCATGTC	223
Db	42	ThrCysLysHisGlnCysGlnGlnGlnGlnTyrThrGluGlyAspLysArgValCys	61
QY	224	GAGTCTCCTCGACCAAGCTCGATGATGATCCTCGTTGTGTCTATGATCTCGAGGACAC	283
Db	62	LeuGlnSerCys-----	65
QY	284	ACTGGCACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACGTTGGCGGCCAACCCCGA	343
Db	65	-----	65
QY	344	GACTACGATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGGACCACT	403
Db	66	-----AspArgTyrHisArgMetLysGlnGlu-----	74
QY	404	GGACGAGGAGCGTGAAGAGAGAGACTGGAGACAACCAAGAGAGAGATTGGAGGCGA	463
Db	75	-----ArgLysLysGlnIleGlnGlnGlnGlnGlnTyrArgGluLysLysGluGlu---	89
QY	464	CAAAGTCATCAGCAGCCAGGAAATAAGGCCCAAGGAGGAGGAGGAGGAGGAGTGG	523
Db	90	-----GluSerArgGluArgGluGluGlu---	97
QY	524	GGAAACACGAGTAGCCATGTGAGGGAAGAAACATCTCGAAACAACCCCTTCTACTTC---	580
Db	98	-----GlnGlnGlnHisGlnGluGlnAspGluAsnProTyrIlePheGlu	113
QY	581	CGTCAAGCGGTTTAGCACCGCTACGGGAACCAACGCTAGGATCCGGCTCTCGAC	640
Db	114	GluAspLysAspPheGluThrArgValGluThrGluGlyGlyArgIleArgValLeuLys	133
QY	641	AGGTTTGACCAAGGTCGAAGGCAGTTTCCAGAATCTCCAGAATCACCGATTATTGTGCAGATC	700
Db	134	LysPheThrGluLysSerLysLeuGlnGlyIleGluAsnPheArgLeuAlaIleLeu	153
QY	701	GAGCCAAACCTAACACTCTTTGTTTCTCCCAAGCACGCTGATGCTGATAACATCCTTGT	760
Db	154	GluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluValValPhePhe	173
QY	761	ATCCAGCAAGGCGAACCCCGTACCGTAGCAATGGCAATGCAATAACAGAAAGCTTTAAT	820
Db	174	AsnIleLysGlyArgAlaValLeuGlyLeuValSerGluSerGluThrGluLysIleThr	193
QY	821	CTTGACGAGGCGCTGCTCAGATCCATCCCGCTTTCATTCCTACATCTTGAACCGC	880
Db	194	LeuGluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArg	213
QY	881	CATGACAAACCAACCTTCAGATAGTAAATCTCCATGCC-----GTTAACAACACCC	934
Db	214	AspGluAsnAspLysLeuPheLeuAlaMetLeuHisIleProValSerValSerThrPro	233
QY	935	GGCCAGTTTGAGGATTTCTCCCGCGAGCAGCCAGACCAATCATCTCTACTTTCAGGGC	994
Db	234	GlyLysPheGluGluPheAlaProGlyGlyArgAspProGluSerValLeuSerAla	253
QY	995	TTGACGAGAAATAGCTTGGAGCGCGCTTCAATCGGAAATTCATAGATGAGGAGGCTG	1054
Db	254	PheSerTrpAsnValLeuGlnAlaAlaLeuGlnThrProLysGlyLysLeuGluAsnVal	273
QY	1055	CTGTTAGAAGAAGATGACAGGAGGTGAGCAAGAGGAGAGGCGGAGCGGATGGAGTACT	1114

```
Db      274 Phe----- 274
QY      1115 CGAGTAGTGAGAACATGAAGCGAGTGATAGTCAAAAGTGTCAAAGGAGCAGCGTTGAAGAA 1174
Db      275 -----AspGlnGlnAsnGlnGlySerIlePheArgIleSerArgGlnValArgAla 292
QY      1175 CTTACTAAGCAGCGCTAAATCCGCTCAAGAAAGGCTCC-----GAAGAAGGGGAGAT 1228
Db      293 Leu-----AlaProThrLysLysSerSerTrpTrpProPheGlyGlyGlu 307
QY      1229 ATCAACCAACCAATCAACTTGAGAGAAGCGGAGCGGATCTTTCTAACTTTGGGAAG 1288
Db      308 SerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGlyTyrGlyArg 327
QY      1289 TTATTTGAGGTGAAGCCA---GACAAGAAGAACCCCGAGCTTCAGACCTGGACATGATG 1345
Db      328 LeuThrGluValGlyProAspAspGluLysSerTrpLeuGlnArgLeuAsnLeuMet 347
QY      1346 CTCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCACTCAAAAGGCC 1405
Db      348 LeuThrPheThrAsnIleThrGlnArgSerMetSerThrIleHisTyrAsnSerHisAla 367
QY      1406 ATGTTTATCGTCTGCTCAACAAGGAACCTGGAAACCTTGAACCTCGTGGCTGTAAAGAAA 1465
Db      368 ThrLysIleAlaLeuValIleAspGlyArgGlyHisLeuGlnIleSerCysProHisMet 387
QY      1466 GAGCAACAACAGAGGGGACCGCGGGAAGAGAGGAGGACGAGACGAAAGACGAAAGAGGAGGGA 1525
Db      388 SerSerArgSerSerHisSerLysHisAspLys----- 398
QY      1526 AGTAACAGAGAGGTGGTAGGTACACAGCGAGGTGTAAGAGGCGGATGTGTTTCATG 1585
Db      399 SerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPheValVal 418
QY      1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTT-----GGC 1639
Db      419 ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuMetIleCys 438
QY      1640 TTCGGTATCAACGCT 1654
Db      439 PheGluValAsnAla 443
```

Search completed: August 24, 2005, 10:09:09
Job time : 62.1903 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 10:05:17 ; Search time 131.012 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3562

Sequence: 1 ataatacatatatattcatc.....cgtttgctggtgtttctcc 2032

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US10728323/runat_23082005_124357_29323/app_query.fasta_1.4757
-DB=Published Applications AA -OEMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -NODES=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10728323 @CGN 1 1 221 @runat 23082005 124357_29323
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3286	92.3	626	10	US-09-847-208-28
2	3286	92.3	626	14	US-10-228-806-2
3	3286	92.3	626	15	US-10-100-303A-7
4	3286	92.3	626	15	US-10-245-871-10
5	3286	92.3	626	15	US-10-253-286-10
6	3286	92.3	626	16	US-10-809-689-95
7	3286	92.3	626	17	US-10-899-551-2
8	3192	89.6	635	17	US-10-899-551-53
9	3052	85.4	634	9	US-09-731-221-78
10	3041	85.4	614	9	US-09-331-631A-21
11	3041	85.4	614	10	US-09-847-208-27
12	3041	85.4	614	14	US-10-147-095-21
13	3041	85.4	614	15	US-10-100-303A-8
14	1394	39.1	268	15	US-10-100-303A-55
15	1282	36.0	605	15	US-10-424-599-153195
16	1281	36.0	605	9	US-09-331-631A-25
17	1281	36.0	605	14	US-10-147-095-25
18	1275.5	35.8	623	15	US-10-424-599-153206
19	1274.5	35.8	623	16	US-10-739-930-8840
20	1268	35.6	605	15	US-10-100-303A-110
21	1201.5	33.7	584	15	US-10-424-599-260105
22	1157	32.5	417	15	US-10-245-227B-14
23	1157	32.5	425	15	US-10-245-227B-12
24	1156	32.5	439	15	US-10-245-227B-1
25	1117.5	31.4	390	15	US-10-245-227B-2
26	874	24.5	324	15	US-10-425-114-44408
27	867	24.3	344	15	US-10-425-114-51703
28	865.5	24.3	666	9	US-09-331-631A-3
29	865.5	24.3	666	14	US-10-147-095-3
30	863.5	24.2	625	9	US-09-331-631A-5
31	863.5	24.2	625	14	US-10-147-095-5
32	849.5	23.8	666	9	US-09-331-631A-1
33	849.5	23.8	666	14	US-10-147-095-1
34	828	23.2	316	15	US-10-425-114-43905
35	828	23.2	316	15	US-10-425-114-49262
36	827	23.2	315	15	US-10-425-114-43971
37	816	22.9	301	15	US-10-425-114-49373
38	791	22.2	590	9	US-09-331-631A-8
39	791	22.2	590	14	US-10-147-095-8
40	790	22.2	300	15	US-10-425-114-43597
41	788	22.1	296	15	US-10-425-114-43984
42	770	21.6	291	15	US-10-425-114-51411
43	761	21.4	149	15	US-10-100-303A-54
44	737.5	20.7	525	9	US-09-331-631A-7
45	737.5	20.7	525	14	US-10-147-095-7

ALIGNMENTS

RESULT 1
US-09-847-208-28
; Sequence 28, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-28

Alignment Scores:				
Pred. No.:	2,4e-261	Length:	626	
Score:	3286.00	Matches:	626	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	92.25%	Indels:	0	
DB:	10	Gaps:	0	
US-10-728-323-1 (1-2032) x US-09-847-208-28 (1-626)				
QY	50	ATGAGAGGAGGGTTTCTCCATGTAGTCTGTTGCTAGGATCCTGTCTGCTTCAGTT	109	
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20	
QY	110	TCTGCAACGCATCCCAAGTCATCATTACCAAGAGAAACAGAGAACCCCTCCGCCAG	169	
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40	
QY	170	AGGTGCTCCAGATTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCGCAGTCT	229	
Db	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60	
QY	230	CGCTGACCAAGCTCAGTATGATCTCTTGTGTATGATCTCTCGAGGACACACTGGC	289	
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80	
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGCTGGCGCGCCACCCGGAGCTAC	349	
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100	
QY	350	GATGATGACCGCGCTCAACCCCGAAGAGAGGAGCGCGATGGGGACCACTGGACCG	409	
Db	101	AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro	120	
QY	410	AGGGAGCTGGAAGAGAGAAGACTTGGAGACAACCAAGAGAAGATGGAGGCGCAAGT	469	
Db	121	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer	140	
QY	470	CATCAGCAGCCAGAAATAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529	
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	160	
QY	530	CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGCTCAAG	589	
Db	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180	
QY	590	CGGTTTAGCCCGCTACGGGAACCAAAACGGTAGATCCGGGTCTCTGAGAGGTTTGAC	649	
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200	
QY	650	CAAAGGTCAAGCAGTTTCAGAAATCCAGAAATCCAGAAATCCAGAAATCCAGAAAT	709	
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220	
QY	710	CCTAACTCTTGTCTTCCCAAGCAGCTGATGATGATGATGATGATGATGATGATGAT	769	
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240	
QY	770	GGGCAAGCCCGTACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG	829	
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260	
QY	830	GGCCATGCACTCAGAAATCCATCCGGTTTCATTTCTCATCTTCAACCGCCATCACAAC	889	
Db	261	GlyHisAlaLeuArgGlyLeuProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280	
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCGCCGTTTAAACACACCCCGGCGAGTTTCAG	949	
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300	
QY	950	TTCCTCCGGCGAGCGGAGACCAATCATCTTCTGAGGCTTCAGCGGAGTAACG	1009	

Db	301	PhePheProAlaLysSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGGAGGCGGCTTCAATGCGGAATTCATGAGATACGAGGGTCTGTGTACAGAGAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn	340
QY	1070	GCAGGAGTGTGACCAAGAGAGAGAGGCGGATCGAGTACTCGGAGTAGTAGAGAAC	1129
Db	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn	360
QY	1130	AATGAAGGAGTATAGTCAAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACG	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGCTCAAAAGAAAGGCTCCGAAGAAGAGGAGATATACCAACCCATCACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGAGCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAG	1309
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAGAACCCCGACTTCAGGACCTGACATGATGCTCACTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu	440
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAACCTTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGAG	1489
Db	461	GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAG	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGlu	500
QY	1550	ACAGCGAGTTGAAGAGGCGATGTGTTCATCATGCCAGAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAla	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTGGCTTCGGTTCAGTCAACGCTGAAAAACAC	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHis	540
QY	1670	ATCTTCTCTGAGGTATAGCAATGTGTATAGACCATAGACAGAGAGAGAGAGAG	1729
Db	541	IlePheLeuAlaGlyAspLysAspValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTGCGTGGGTGAACAAGTTCAGAAAGCTCATCAAAACCCAGAGGAAT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTTGTAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTCTGAGAAAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCTGTAGAAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1909
Db	601	SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGlu	620
QY	1910	ATTTTGAAGGCTTTTAAAC 1927	
Db	621	IleLeuLysAlaPheAsn 626	

RESULT 2

US-10-228-806-2
; Sequence 2, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions

```
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228, 806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-2

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 14 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-228-806-2 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGCTAGGGATCCTTGTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TGTGCAACGATGCGCAAGTCATCACCTTACCAAGAAACAGAGAACCCCTGGGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGATGTTCAACAGGAACCGGATGCTTCAAGCAAAAGCATGCCAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCCAGTATGATCTCGTTGTGTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACAGTTCCCTCAGGGAGCGGACACAGTGGCGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGAGAGAGGAGGAGCGGATGGGGACCACTGGACCG 409
DB 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrProAlaGlyPro 120
QY 410 AGCGAGCGTGAAGAGAAGAACTCGACACCAACCAAGAAAGATTGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGACCCACGGAATAAAGCCCGAAGAGAGAGAAACAGAGTGGGGAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCGAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CTTAACACTCTTGTCTTCCAGAGCAGCTGATGCTGATAAATCTTGTATTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGAACCGTAGCAATTAACAGAAAGAGCTTTTAACTCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
```

```
QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTCTACATCTTGAACCGCCATGACAA 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGCCAGTTTTCAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCGCAGACCAATCATCTTCTACTTGCAGGGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCGCTTCAATGCGGAATTCATATGAGATACGAGGGTGTCTGTAGAGGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGCAGAGCGCATCGAGTACTCGGAGTAGTGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrPheSerThrArgSerGluAsn 360
QY 1130 AATGAAGGAGTGATAGTCAAGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGGCTCCGAAGAAGGAGAGATATCACCACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGCGCGCCGATCTTCTAAACAACCTTGGGAAGTATTATTGAGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAAGACCCCGACCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCTGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GSAACCTGGAACCTTGAACCTCGTGGCTGTGAAGAAAGAGCAACACAGAGGGGCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 CAAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGTGGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGGAAGCGGATGTGTTTCATCATGCGCAGCAGCTCATCCAGTCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTCGCTTCCGCTTCAAGCTCAAGCTGAAACCAACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTGAGGTGATAAGGACAAATGTGTAGACACCATAGAGAAGCAAGCGAAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAAACAGAAAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCGCTCCTCAATCTCAATCTCAATCTCGCTCGCTCCTCCTCAGAAAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyGlyGlyGlyGlyGlyGly 620
```

```
QY 1910 ATTTGAAGCGTTTAAAC 1927
Db |||||||
621 IleLeuLysAlaPheAsn 626

RESULT 3
US-10-100-303A-7
; Sequence 7, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; TYPE: PRT
; LENGTH: 626
; ORGANISM: Atachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-7

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-100-303A-7 (1-626)
QY 50 ATGAGAGGAGGAGTTCCTCACTGATGCTGTGTAGGGATCCTTGCTCGTTCAGTT 109
Db |||||||
1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCAAGTCATCACCCTTACCAAGAAAGAAACAGAGAACCCCTGCGCCCGAG 169
Db |||||||
21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCGCTCCAGATTGTCACAGAGACCGGATGATCTGAAGCAAAAGCATCGAGTCT 229
Db |||||||
41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTGCTGTGCTATGATCCTCGAGGACACACTGCG 289
Db |||||||
61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCCGAGACTAC 349
Db |||||||
81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGAGCGCGATGGGGACCACTGGACCG 409
Db |||||||
101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGGAGCGTGAAGAGAGAAGACTGGAGACAAACCAAGAGAGATTGGAGGCGCAAGT 469
Db |||||||
121 ArgGluArgGluArgGluArgGluArgGlnProArgGlnProArgGlnProArgGlnPro 140
QY 470 CATCAGCAGCCACGGAAATAAGGCCGCGAAGGAAGAGAGAGAAACAAGAGTGGGGAACA 529
Db |||||||
141 HisGlnGlnProArgLysIleArgProGluArgGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CCAGTAGCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTCCCGCTCAAGG 589
Db |||||||
161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCCGCTACGGGAACCAAAACGCTAGGATCCGGGTCTCGCAGAGGTTTGAC 649
Db |||||||
181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

650 CAAAGGTCAAGCAGTCTTTCAGAAATCCAGAAATCACCGTATTTGTGCAGATCAGGCCAAA 709
Db |||||||
201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCAGTGTGATGATACATCCTTGTATCCAGCNA 769
Db |||||||
221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCAGCCTAGCAATGCAATTAACAGAAAGAGCTTTAATCTTCACGAG 829
Db |||||||
241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GCCCATGCACCTCAGAAATCCCATCCGTTTTCATCTTCTACATCTTTGAACCGCATCACAAC 889
Db |||||||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGCAGTTTTCAGGAT 949
Db |||||||
281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCAGCAGACCAATCATCTTCTACTTGCAGGGCTTCAGCAGGAAATACG 1009
Db |||||||
301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATGCCGAATTCATGAGATACCGAGGGTCTCTTTAGAAAGAGAT 1069
Db |||||||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY 1070 GCAGAGGTGACCAAGAGCAGAGAGAGGCGAGGCGATCGAGTACTCGGAGTAGTCAGAAC 1129
Db |||||||
341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGGTGATAGTCAAAAGTGTCAAAGGAGACAGCTTGAAGAACTTTACTAAGCACGCT 1189
Db |||||||
361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAAGAGGAGAGATATCAACAACCCCAATCAACTTG 1249
Db |||||||
381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGCGGAGCCCGATCTTTCTAACAACTTGGGAAGTTATTTGAGGTGAGGACGAGAC 1309
Db |||||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db |||||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCTCGTCAACAAA 1429
Db |||||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGTGAACCTTGAACCTCGTGGCTGTAGAAGAGCAACACACAGAGGGAGCGCGG 1489
Db |||||||
461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGAGCAAGCAAGAGAGGAGGAGTAACAGAGAGCTCGTAGGTAC 1549
Db |||||||
481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTGAAGAAAGCGATGTGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATC 1609
Db |||||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCGAACTCCATCTGCTGGCTTGGTATCAACGCTTGAAGAAACCAACACAGA 1669
Db |||||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTGCAGGTGATAAGCAATGTGATAGACCATAGAGAGAGCAAGCAGCGAAGGAT 1729
Db |||||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGAAATCT 1789
```



```
Db 481 GluGluGluGluAspGluGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGAACGGATGTTTCATCATGCCAGCAGCTCATCCAGTACCATC 1609
Db 501 ThrAlaArgLeuGlyGluGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY 1610 AACGCTTCCTCCGAATCCCATCTGCTTGCTTCCGTTATCAACGCTGAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCCTTCAGGTGATGAAGACATGTGATAGACCAGATGAGAACGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspIleAspValIleAspGlnIleGluGlyGlnAlaIleAsp 560
QY 1730 TTAGCATTCCTCGGTCGAGTGAACAAAGTTGAGAGCTCATCAAAACACAGAAAGTCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluGlyLeuIleAsnGlnIleGlySer 580
QY 1790 CACTTTGTGAGTGCTGCTCCTCAATCTCAATCTCCTGCTCCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluGlyGlu 600
QY 1850 TCTCTGAGAAACAGATCAAGAGAGGAGAAACCAAGAGGAGGAGGTCCTCTTCA 1909
Db 601 SerProGluGlyGluAspGlnGluGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTAAAC 1927
Db 621 IleLeuGlyAlaPheAsn 626

RESULT 5
US-10-253-286-10
; Sequence 10, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Atachis hypogaea
US-10-253-286-10

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-253-286-10 (1-626)
QY 50 ATGAGAGGAGGTTTCCCATGTGCTGCTAGGATCCTTGCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCCAAGTCATCACCTTACCAGAGAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaIleAspSerSerProTyrGlnIleGlySerThrGluAsnProCysAlaGln 40
QY 170 AGTGCGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
```

```
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuGlyGlnLysAlaCysGluSer 60
QY 230 CGCTCACCACGCTCGAGTATGATCCTCGTGTGCTATGATCCTCGAGGACACACTCGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGCGGACACGCTGGCCGCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGAAAGAGAGAGAGAGCCGATGGGACCACTCGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgGlyArgGlnProAlaGlyPro 120
QY 410 AGGAGCGCTGAAGAGAGAGAGAGCTGGAGAGACCAACAGAGAGATTGGAGGCGCAAGT 469
Db 121 ArgGluArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGCAAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgGlyIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAGAGAAACATCTCGGAACAACCTTCTACTTCCCGCTCAAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGNAACCAAAACGTTAGGATCCGGGTCTCGAGAGGTTTGCAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGluArgPheAsp 200
QY 650 CAAAGGTCACAGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCAGAGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCAGCGTAGCAATGCGCAATACAGAAAGAGCTTTAACTTTCACCGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACCTCAGAATCCCATCCGTTTCTATTTCTCATCTTGAACCGCCATCAGCAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTGAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCGAGACCAATCATCTTCTGAGGCGTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGCGCGCTTCAATGCGGAATTCATAGATACATCGAGGGTGCTCTTGAAGAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGAGGTGAGCAAGAGAGAGAGGCGAGGCGATGAGTACTCGGAGTAGTAGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATGATCAAGTGTCAAAGGAGACGTTGAAGAACTTACTTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAAAGAGGCTCCGAAAGAGAGGAGATATCACCACACCCCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGCGGAGCCCGATCTTTCTAAACACTTGGGAAGTTATTGTAGGTGAGCGACAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
```

QY 1310 AAGAAGACCCAGCTTCAGGACCTGGACATGCTCACCCTGTGTAGAGATCAAGAA 1369
Db 421 LysLeuAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCGTCGTCACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAlaLys 460
QY 1430 GGAATCGGAACCTTGAATCTCGGTGTAAGAAAAGACAAACAGAGGGGACGGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
QY 1490 GAAGAAGGAGACGACGACGAGAGAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGlu 500
QY 1550 ACAGCCAGGTGAAGAGGCGATGTGTCATCATGCGCAGCATCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCCTCGAATCCATCTCGTTCGCTTCGCTATCAAGCTCAAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTTCAGGTGATAGGACATGTGATAGACAGATAGAGCAAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTCGGTGCGGTGAACCAAGTTGAGAAGCTCATCAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCCGTCGTCCTCGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGTTCACCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 6

US-10-809-689-95
; Sequence 95, Application US/10809689
; Publication No. US20040265342A1
; GENERAL INFORMATION:
; APPLICANT: Eric Potter Clarkson
; TITLE OF INVENTION: Methods and compositions for desensitisation
; FILE REFERENCE: 5538/1010
; CURRENT APPLICATION NUMBER: US/10/809,689
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: PCT/GB99/00080
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: GB/9800445.0
; PRIOR FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: GB/9820474.6
; PRIOR FILING DATE: 1998-09-21
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 95
; LENGTH: 626
; TYPE: PRT
; ORGANISM: *Arachis hypogaea*
US-10-809-689-95
Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 16 Gaps: 0
US-10-728-323-1 (1-2032) x US-10-809-689-95 (1-626)
QY 50 ATGAGAGGGAGGGTTTCTCCACTGATGCTGTGTAGGGATCCTGTGCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCGCAAGTCATCACCTTACACAGAGAAAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCTCCAGAGTTGTCAACAGAGAACCGGATAGCTTGAAGCAAAAGGATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTGTCTATGATCCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACAGTGGCGCCCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGAGGAGCGCGATGGGACCACTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGAGCGTGAAGAGAGAGAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCAACT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCATGACGCCACGGAATAAGGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160
QY 530 CGAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTTACCTCCGCTCAAGG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGCTAGGATCCGGGTCTCGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCGCAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTCTTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCACCGTAGCAATGCGCATACAGAGAGCTTTTAACTTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCAATTTCTTACATCTTGAACCGCATGACAAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGCTTAACACACCCGCGCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCCTCCGCGGAGGAGCGGAGACCAATCATCTTCTACTTTCAGGCTTCAGCAGGATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCGCTTCAATGCGGAAATTCAATGAGATACGAGGGGTGCTGTTAGAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340


```
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrlleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGTAAATCTCCATCCCGTAAACACACCCCGCCAGTTTGAGGAT 949
Db 281 GluAsnLeuArgValAlaIysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCAGACCATCATCTCTACTTGCAGGGCTTCAGCAGGAATCG 1009
Db 301 PhePheProAlaSerArgAspGlnSerSerTyrlleGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGCGAATCAATCAGATACGGAGGCTGCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValleuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGAGCGGATGAGGATCTCGAGTACTGAGAAC 1129
Db 341 AlaGlyGlyGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAGGTGTCAAGGACGCTTGAAGACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValIysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGGCTCCGAGAGAGGAGATATCAACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGAGCCGATCTTTCTAACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGACCCCGAGCTTCAGGACCTGGACATGCTCACCTGCTGATAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGTCCACACTTCAACTCAAGGCCATGTTATCTCGTGGTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGGTGTGAAGAAAGAGCAACACAGAGGGAGCGCG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGACCAAGACGAGAGAGGAGGAGGAGTAACAGAGAGTGCCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTGAAGGAAGCGATGTGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY 1610 AACGCTTCCTCGAATCCCATCTGCTTGGCTTCGGTATCAAGCTGAAACCAACCA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTCAATAGGCAATGTATGATAGACAGATAGAGCAAGCAAGGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTGGGTGGGTGAACAGTTGAGAAGCTCATCAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGGAGGGTCACTCCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
```

```
Db 621 IleLeuLysAlaPheAsn 626
RESULT 8
US-10-899-551-53
; Sequence 53, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 635
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-53
Alignment Scores:
Pred. No.: 1,33e-253 Length: 635
Score: 3192.00 Matches: 605
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.83% Mismatches: 0
Query Match: 89,61% Indels: 0
DB: 17 Gaps: 0
US-10-728-323-1 (1-2032) x US-10-899-551-53 (1-635)
QY 110 TCTGCAACGATCGCCAAAGTCATCACCTTACCAGAGAAGAAAAACAGAGAACCCCTGCGCCAG 169
Db 17 SerSerThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 36
QY 170 AGTGCTCTCCAGAGTTGTCAACAGAGAACCGGATGACTTGAAGACAAAGCATGCGAGTCT 229
Db 37 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 56
QY 230 CCCTCCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGAGCACACTGCG 289
Db 57 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 76
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACGTCGCGCCGCAACCCCGAGAGACTAC 349
Db 77 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 96
QY 350 GATGATGACCGCCGTCACCCCGAAGAGAGAGGAGGCGCGATGGGACCACTGGACCG 409
Db 97 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 116
QY 410 AGGGAGCTGGAAGAGAGAGACTGGACACCAACAGAGAGATTTGGAGGCGACCAACT 469
Db 117 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 136
QY 470 CATCAGACGCCACGAAATAAGGCCCGAAGGAAGAGAGAGAGAAACAGAGTGGGGAAACA 529
Db 137 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 156
QY 530 CCAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTCCGTCAGG 589
Db 157 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 176
QY 590 CGGTTTAGCCCGCTACGGNACCAAAACGCTAGGATCCGGTCTCTGCGAGGTTTGAC 649
Db 177 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 196
QY 650 CAAAGGTCAAGGCACTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCGCAA 709
Db 197 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 216
```

```
QY 710 CCTAACACTCTTGTCTTCTCCAGCAGCTGATGCTGATAAATCATCTCTTGTATCCAGCAA 769
Db 217 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln 236
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATATGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db 237 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 256
QY 830 GGCATGCACTCAGATCCATCCGCTTTCATTTCTTACATCTTGAACCGCATCACAC 889
Db 257 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 276
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGCTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 277 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 296
QY 950 TTCTTCCCGGAGCAGCAGCAGCAATCATCTCTTCTCAGGGCTTCAGCAGGAATACG 1009
Db 297 PhePheProAlaSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 316
QY 1010 TTGGAGCGCCTTCAATCGGNAATTCATATGATACGAGGCTGCTGTAGAGAGAT 1069
Db 317 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 336
QY 1070 GCAGGAGGTGACAGAGAGAGAGCGCAGATGAGTACTCGAGTAGTGAGAAC 1129
Db 337 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 356
QY 1130 ATGAAGGAGTAGATCAAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCT 1189
Db 357 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 376
QY 1190 AATCCGCTCTCAAGAAAGCGCTCCGAGAGAGGAGGATATCACCAACCAATCACTTG 1249
Db 377 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 396
QY 1250 AGAGAAGCGCAGCCCGATCTTCTTCAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 397 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 416
QY 1310 AGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACTCTGTGTAGATCAAGAA 1369
Db 417 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 436
QY 1370 GGAGCTTTCATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCTGCTCAACAAA 1429
Db 437 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 456
QY 1430 GGAACGTGAAACCTTGAACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGCGCGG 1489
Db 457 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 476
QY 1490 GAAAGAGGAGAGCAAGACGAGAGAGAGGAGGAGTAACAGAGAGTGGTGGTAGTAC 1549
Db 477 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 496
QY 1550 ACAGCGAGGTTCAGGAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTACCATC 1609
Db 497 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 516
QY 1610 AACCGTTCCTCCGAACTCCATCTGCTTGGCTTTCGGTATCAACGCTCAAAACCAACACAGA 1669
Db 517 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 536
QY 1670 ATCTTCTTCCAGGTGATTAAGACAAATGTGATAGACAGATAGAGAGCAAGCAGGAT 1729
Db 537 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 556
QY 1730 TTAGCAATCCCTGGGTCCAAAGTTGAGAGCTCATCAAAAACAGAAAGGAATCT 1789
Db 557 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 576
```

```
QY 1790 CACTTTGTAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCT 1849
Db 577 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 596
QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGAGGAGGAGGCTCCACTCTTCA 1909
Db 597 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 616
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 617 IleLeuLysAlaPheAsn 622
RESULT 9
US-09-731-221-78
; Sequence 78, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 78
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-78
Alignment Scores:
Pred. No.: 4,47e-242 Length: 634
Score: 3052.00 Matches: 584
Percent Similarity: 96.86% Conservative: 3
Best Local Similarity: 96.37% Mismatches: 19
Query Match: 85.68% Indels: 0
DB: 9 Gaps: 0
US-10-728-323-1 (1-2032) x US-09-731-221-78 (1-634)
QY 110 TCTGCAACGCATGCCAGTCATCACTTACCAGAGAAACACAGAGAACCCCTGCGCCAG 169
Db 16 SerSerThrHisAlaLysSerSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 35
QY 170 AGTGCTCTCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 36 ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer 55
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTATGATCCTCGAGGACACACTCGC 289
Db 56 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 75
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGACACGTCGCCGCCCAACCCCGGAGCTAC 349
Db 76 ThrThrAsnGlnArgSerProProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 95
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGAGAGCCCGATGGGGACAGCTGGACCG 409
Db 96 AspAspAlaArgArgGlnProArgAlaGluGluGlyArgTyrGlyProAlaGlyPro 115
QY 410 AGGGAGCGCTGAAAGAGAGAGAGAGAGCTGGAGACACCAAGAGAGATTTGGAGCGCAAGT 469
Db 116 ArgGluArgGluArgGluGluAspAlaArgGlnProArgGluAspTyrPalArgProSer 135
QY 470 CATCAGCAGCCACGAAATAAGGCCCGCCAGAGAGAGAGAGAGAACAGAGTGGGGAACA 529
Db 136 HisGlnGlnProArgLysAlaArgProGluGluGluGluGluGluGluGluGluGluThr 155
QY 530 CCAGTAGCCATGTGAGGAGAGAGAACATCTCGGAACAACCCCTTCTTCTTCTCCGTCAGG 589
```

156 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 175
QY 590 CGGTTTAGACCCCGTAGCGGAACCAAAACGGTAGGATCCGGTCTCGAGAGTTTGAC 649
Db 176 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGluArgPheAsp 195
QY 650 CAAAGCTCAAGGCAGTTTCAGATCTCCAGATCCAGATCACCGTATTCTGCAGATCGAGCCAA 709
Db 196 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 215
QY 710 CCTAACACACTCTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAA 769
Db 216 ProAsnThrLeuValLeuProLysHisAlaAspAlaAsnIleLeuValIleGlnGln 235
QY 770 GGCACAGCCACCGTAGCCGTAGCAATGCGAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 236 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 255
QY 830 GGCATGCACTCAGATCCCATCCGCTTTCATTTCTTACATCTTGAAACGCCATGACAC 889
Db 256 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 275
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 276 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnMetGluAsp 295
QY 950 TTCTTCCCGGAGCAGCCGAGACCAATCATCTTCTGAGGGCTTCAGAGGAATACG 1009
Db 296 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheAlaArgAsnThr 315
QY 1010 TTGAGGCGCCTTCAATCGGNAATCAATGATAGATACGAGGGTGTCTGTAGAGGAT 1069
Db 316 LeuGluAlaAlaPheAsnAlaGluAlaAsnGluIleArgArgValLeuLeuGluGluAsn 335
QY 1070 GCAGGAGGTGAGCAAGCAGGAGCGCAGAGCGGATGAGTACTCGAGTAGTGTAGAAC 1129
Db 336 AlaGlyGlyGluGlnGluAlaArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 355
QY 1130 AATGAAGAGTAGTATGATCAAGTGTCAAGAGGACACGTTGAAAGACTTACTAAGCACGT 1189
Db 356 AsnGluGlyValIleValIleValIleValSerLysGluHisValGluGluLeuThrLysHisAla 375
QY 1190 AATCCGTCTCAAGAAAGCGCTCCGAGAGAGGGAGATATCACCAACCAATCACTTG 1249
Db 376 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProAlaAsnLeu 395
QY 1250 AGAAGAGCGAGCCCGATCTTTCTAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 396 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuAlaGluValLysProAsp 415
QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 416 LysLysAsnProGlnLeuGluAspLeuAspMetMetLeuThrCysValGluIleLysGlu 435
QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCGTCTGCTCAACAA 1429
Db 436 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 455
QY 1430 GGAACCTGGAACCTTCAACTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGCGCGG 1489
Db 456 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 475
QY 1490 GAAGAAAGAGGAGGACCAAGACCAAGAGAGGGAGGAAGTAAACAGAGAGTGGTGGTAGTAC 1549
Db 476 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgAlaTyr 495
QY 1550 ACAGCAGGTTGAAGAAAGCGATGTTCATCATCGCAGCAGCTCATCCAGTAGGCATC 1609
Db 496 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 515
QY 1610 AACGCTTCTCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACACAGA 1669
Db 516 AsnAlaSerSerGluLeuAlaLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 535

1670 ATCTTCTCTTGCAGGTGATTAAGGACAAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGAT 1729
Db 536 IlePheLeuAlaGlyAspAlaAspAsnValIleAspGlnIleGluLysGlnAlaLysAla 555
QY 1730 TTAGCATTTCCCTGGGTCCGGTCAACAAGTTGAGAAGCTCATCAAAAACCCAGAGGAATCT 1789
Db 556 LeuAlaAlaProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 575
QY 1790 CACTTTGTGAGTGCCTCGTCTCTCAATCTCAATCTCTCGTCTCTCTCAGAAAGAG 1849
Db 576 HisPheValAlaAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 595
QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGAGGAGGAGGTCCTCTCTTCA 1909
Db 596 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 615
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 616 IleLeuLysAlaPheAsn 621
RESULT 10
US-09-331-631A-21
; Sequence 21, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULLIN23.001APC
; CURRENT APPLICATION NUMBER: US/09/331,631A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
US-09-331-631A-21
Alignment Scores:
Pred. No.: 3,55e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 85.37% Indels: 18
DB: 9 Gaps: 7
US-10-728-323-1 (1-2032) x US-09-331-631A-21 (1-614)
QY 50 ATGAGAGGAGGGTTCCTCACTGATGCTGTGTAGGATCCTTGTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGACGATGCCCAAGTCATCATTACCCGAGAAAAACAGAGAACCCCTCGCCGAC 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGTCCCTCCAGAGTTGTCAACAGCAACCGGATGACTTGAAGCAAAAGCATCCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGCG 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTGGCCGCCAACCCCGAGACTAC 349

Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGCTCAACCCGGAAGAGAGGAGCGCGATGGGGACCGACGTCGACCG 409
Db 95 AspAspArgArgGlnProArgArgGluGluGlyArgTyrProAlaGluPro 114
QY 410 AGGAGCGTCAAGAGAAGAAGACTGGAGACAACCAAGAGAAGATTTGGAGGGACCAAGT 469
Db 115 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 134
QY 470 CATCAGCAGCCGGAATAATAGCCCGGAAGAGAGAGAGAGACAAGAGTGGGGAGACA 529
Db 135 HisGlnGlnProArgGlySerIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 154
QY 530 CAGGTTAGCGTGTGAGGAAGAACAATCTCGGAACAACCCCTTCTACTTCCCGTCAGG 589
Db 155 ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGGGTAGGATCCGGGTCTCGCAGAGGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCCAAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CTTAACACTCTTGTCTCCCAAGCAGCGTGTGATAACATCTCTGTTATTCAGACAA 769
Db 215 ProAsnThrLeuValLeuProLysHisIleAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTGCAGTCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCCATGCACTCAGATCCCATCCGTTTCATTTCTACATCTTGAAACCGCATCAGCAAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGAGTAGCTAAATCTCATGCCGCTTAAACACACCCGCCAGTTTCAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGCGAGCGGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATAGC 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGGAGGCGCCTTCATCCGGAATTCATAGATACGAGGGTCTGTAGAGAGAAAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGGTGACAGAGGAGAGGCGGAGCGGATGGAGTACTCGGAGTAGTCAGAAC 1129
Db 335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTAGTAGTCAAAAGTGTCAAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTCTCAAGAAAGGCTCCGAGAGAGGAGATATCACCAACCAATCAACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAGAAGGCGGCGGATCTTTCTAACACTTTGGAGAGTATTATTTGAGGTGAAGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLysSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGACTTCAGGACCTGGACATGATCTCACTCTGTGTAGAGATCAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCGTCGTCGTCAACAA 1429

Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 452
QY 1430 CGAACTCGAAACCTTGAACCTCGTGGCTGTAGAAAGAGCAACAACAGAGGGGACGGCGG 1489
Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGGAGGACGAAGCAAGAGAGGAGGGAAGTAACAGAGAGGTG 1540
Db 473 GluGlnGluTyrGluGluGluGluAspGluGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTTGAAGGAAGCGATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCCTCCGAACCTCATCTGCTGGTTCGGGTATCAACCGCTCAAAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAAGAAATCTTCTTCAGGTGATAGGACAAATGTGATAGACAGATAGAGAACAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 CGGAAGGATTTAGCAATCCCTGGGTCGGGTGAAACAGTTGAGAACCTCATCAAAACCAAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGATCTCTGAGAAAGAGATCAAGAGGAGGAAAAACCAAGGAGGAGGAGGTCCA 1900
Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTTCAATTTCAAGCGCTTTAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614
RESULT 11
US-09-847-208-27
; Sequence 27, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: US67,002A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-27
Alignment Scores:
Pred. No.: 3.55e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 85.37% Indels: 18
Dbs: 10 Gaps: 7
US-10-728-323-1 (1-2032) x US-09-847-208-27 (1-614)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGTGTGTAGGATCCTTGCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20

110 TCTGCAACGATCCAGTCACTCACCTTACCAGAAAGAAACAGAACCCCTGCGCCAG 169
Db SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229
Db ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
230 CGTGCACCAAGTCGAGTATGATCTCGTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGGCCGCAACCCGGAGACTAC 349
Db AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
350 GATGATGACCGCGCTCAACCCCAAGAGAGAGGAGCGGATGGGGACACGATGACCG 409
Db AspAspAspArgGlnProArgArgGluGlyArgGlyGlyArgTrpGlyProAlaGluPro 114
410 AGGGAGCGTGAAGAGAGAGAGTGGAGACAAACCAAGAGAGATTGGAGGCGCAAGT 469
Db ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 134
470 CATCAGCAGCACGGAAATAAGSCCGAAGGAAGAGAGGAAACAAGAGTGGGGAACA 529
Db HisGlnGlnProArgLysIleArgProGlyGlyArgGlyGlyGlnGlnGlnTrpGlyThr 154
530 CCAGGTAGCATGTGAGGAAGAAACATCTCGAAACAACCTTCTACTTCCCGTCAAGG 589
Db ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCGGGTCTCGAGAGTTTGAC 649
Db ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
650 CAAAGTCGAAGCAGTTTCAGATCTCCAGATCCAGATTCGTCGAGATCGAGGCCAA 709
Db GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
710 CCTAACACTCTTGTCTTCCAGACACGCTGATGCTGATAACATCTCTGTTATCCAGCAA 769
Db ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
770 GGGCAAGCCACCGTACCGTAGCAATGCAATTAACAGAAAGAGCTTTAATCTTGACGAG 829
Db GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
830 GGCATGCACTCAGATCCCATCCCGTTTCATTTCTCTACATCTTGACCGCCATGACAAC 889
Db GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTAAACACACCCCGCCAGTTTGAGGAT 949
Db GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
950 TTCTTCCCGCGAGCAGCGGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG 1009
Db PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
1010 TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGGTGTGTAGAGAGAT 1069
Db LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
1070 GCAGGAGGTGAGCAAGAGGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
Db AlaGlyGlyGlnGlnGluArgGlyGlnArgArgArgSerThrArgSerSerAsp--- 353
1130 AATGAAGGAGTGTAGTCAAGTGTCAAAGGACACGTTGAAGAATCTTACTAAGCAGCT 1189
Db AsnGluGlyValIleValLysValSerLysGluHisValGlnGlnLeuThrLysHisAla 373
1190 AATCCGCTCTCAAGAAAGCTCCGGAAGAGAGGGAGATATACCAACCCCAATCAACTTG 1249

374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
1250 AGAGAAGCGAGCCCGATCTTTTAACTTTGGGAAGTTATTTTTCAGGTGAAGCCAGAC 1309
Db ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
1310 AAGAAGAACCCCGACCTTCAGGACCTCGACATGATGCTCCTCTGTGTAGATCAAGAA 1369
Db LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTATCTGCTCGTCAACAAA 1429
Db GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
1430 GGAATCGGAACCTTGAACCTCGTGTGTAAGAAAGAGCAACAACAGAGGGAGCGCG 1489
Db GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgGluVal 472
1490 GAA-----GAAGAGGAGACGACGACGAGAGGAGGAGGAGTAAACAGAGAGGTG 1540
Db GluGlnGluTrpGluGluGluGluAspGluGluGluGluGlySerAsnArgGluVal 492
1541 CGTAGGTACACAGCAGGTTGAAGGAGCGGATGTGTTTCATATGCCAGCAGCTCATCCA 1600
Db ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
1601 GTAGCCATCAACGCTTCTCCGAATCCATCTGCTTGGTTCGGTATCAACGCTGAAAC 1660
Db ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
1661 AACCAAGATCTTCTTCAGGTATTAAGGACAATGTGATAGACAGATAGAGAGCAA 1720
Db AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
1721 GCGAAGGATTTAGCATTCCTCGGTGCGGTGAACAGTTGAGAAGCTCATCAAAACCAAG 1780
Db AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGln 572
1781 AAGGAATCTCACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCT 1840
Db ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
1841 GAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGCGGAGGTCCA 1900
Db SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605
1901 CTCTTTCAATTTTGAAGGCTTTTAAAC 1927
Db LeuLeuSerIleLeuLysAlaPheAsn 614
RESULT 12
US-10-147-095-21
; Sequence 21, Application US/10147095
; Publication No. US20030171274A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULLN23.001APC
; CURRENT APPLICATION NUMBER: US/10/147,095
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21


```
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-8

Alignment Scores:
Pred. No.: 3-55e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 85.37% Indels: 18
DB: 15 Gaps: 7

US-10-728-323-1 (1-2032) x US-10-100-303A-8 (1-614)

QY 50 ATGAGAGGGAGGTTTCTCCTACCTGATGCTGTTCTAGGATCCTTGTCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGATCCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGlnAsnProCysAlaGln 38

QY 170 AGTGCTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58

QY 230 CGTGTGACCAAGTCGAGTATGATCTCTGTTGTCTATGATCCTCGAGACACACTGCG 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74

QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgClyArgGlnProGlyAspTyr 94

QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGGAGGCGCGATGGGGACCAAGCTGGACCG 409
Db 95 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114

QY 410 AGGAGCGTGAAGAAGAGACTCGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
Db 115 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer 134

QY 470 CATCAGCCACCGGAATAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 154

QY 530 CAGGTAGCGCATGTAGAGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174

QY 590 CGGTTTACCGCGTACCGGAACCAAAACCGTAGGATCCGGTCTCGAGAGGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194

QY 650 CAAAGTCAGGCAGTTTCAGATCTCCAGATCACCGTATTGTGAGATCGAGGCCAAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214

QY 710 CCTAACACTCTTGTCTCCAGAGCAGCTGATGCTGATAACATCTCTGTTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234

770 GGGCAAGCCACCGTGAACCTTAGCAATGCAANTAGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
830 GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCCTACATCTTGAACCGCCATGACAAC 889
255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAGGAT 949
275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
950 TTTCTCCCGCGAGCAGCCGAGACCAATCATCTTCTTTCAGGGCTTCAGCAGGAATACG 1009
295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
1010 TTGGAGCGCGCTCAATGCGGAATTCATGATGATACGAGGGTGTCTTTAGAGAGAAAT 1069
315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
1070 GCAGGAGGTGAGCAAGAGAGAGAGAGCGGATGAGTACTCGAGTAGTAGAGAAC 1129
335 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
1130 AATGAGGAGTGATAGTCAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTTAAGCACGCT 1189
354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
1190 AATCGCTCTCAAGAAAGGCTCCCAAGAGAGGAGGATATCACCAACCCATCAACTTG 1249
374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
1250 AGAGAGGCGAGCCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGTGAAGCCAGAC 1309
393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
413 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 432
1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGATGTTATCGTCTCGTCAACAAA 1429
433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
1430 GGAACCTGGAACCTTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGAGCGCG 1489
453 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 472
1490 GAA-----GAGAGAGAGAGCGAAGACGAGAGAGGAGGAGGAGTTAACAGAGAGGTG 1540
473 GluGlnGluTyrGluGluGluGluAspGluGluGluGluGlySerAsnArgGluVal 492
1541 CGTAGGTACACAGCGAGTTGAAGAGCGGATGTTTCATCATCGCCAGCAGCTCATCCA 1600
493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
1601 GTAGCCATCAACCGCTTCCTCCGAACTCCATCTCGTTCGTTGGCTTCGATCAACGCTGAAAAC 1660
513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
1661 AACCAAGAAATCTTCTTCAGGTGATAGGACAATGTGATAGACAGATAGAGAGCAA 1720
533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
1721 GCGAAGGATTTAGCATTCCTCGGTCGGGTGAACAGTTGAGAAGCTCATCAAAACCCAG 1780
553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
1781 AAGGAATCTCACTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGCTCCT 1840
573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
```


Qy	1941	GAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAAAAACCAAGGAGGGAGGGTCCA	190
Db	589	-----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro	605
Qy	1901	CTCCTTCCTCAATTTTGAAGGCTTTTAAAC	1927
Db	606	LeuLeuSerIleLeuLysAlaIlePheAsn	614
RESULT 14			
US-10-100-303A-55			
; Sequence 55, Application US/10100303A			
; Publication No. US20030202980A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, et al.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction			
; TITLE OF INVENTION: to Allergy			
; FILE REFERENCE: 2002834-0166			
; CURRENT APPLICATION NUMBER: US/10/100.303A			
; CURRENT FILING DATE: 2002-03-18			
; NUMBER OF SEQ ID NOS: 138			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 55			
; TYPE: PRT			
; LENGTH: 268			
; ORGANISM: Arachis hypogaea, Prot./Nucleo-Ara h 1			
US-10-100-303A-55			
Alignment Scores:			
Pred. No.:		1.04e-105	Length: 268
Score:		1394.00	Matches: 268
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		39.14%	Indels: 0
DB:		15	Gaps: 0
US-10-728-323-1 (1-2032) x US-10-100-303A-55 (1-268)			
Qy	485	AAATATAGGCCCCGAGGAAGAGAGGAGAACAGAGTGGGGAACACAGGTAGCCATGTG	544
Db	1	LysIleArgProGluGlyA-ggLuGlyGluGlnLurpGlyThrProGlySerHisVal	20
Qy	545	AGGAAGAACAATCTCCGGAACAACCCCTTCTACTTCCCGTCAAGGCGGTTTAGCACCCGC	604
Db	21	ArgGluLthrSerArgAsnAsnProPheTyPheProSerArgArgPheSerThrArg	40
Qy	605	TACGGGAACAAACCGTAGGATCCGGGTCTCGCAGAGTTTGACAAAGGTCAAGGCAG	664
Db	41	TyrGlyAsnGlnAsnGlyA-gIleArgValLeuGlnArgPheAspGlnArgSerArgGln	60
Qy	665	TTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACTCTTGT	724
Db	61	PheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLysProAsnThrLeuVal	80
Qy	725	CTTCCCAAGCACGCTGATGTGTATAACATCTTGTATCCAGCAAGGCGCAAGCCCGTG	784
Db	81	LeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGlnGlyGlnAlaThrVal	100
Qy	785	ACGTTAGCAATATGCAATATACAGAAAGAGCTTTTAATCTTCGAGAGGCCATGCACTCAGA	844
Db	101	ThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGluGlyHisAlaLeuArg	120
Qy	845	ATCCCATCCGGTTTCATTTCCTACATCTTGAACCGCCATGACACCAAGACCTCAGAGTA	904
Db	121	IleProSerGlyPheIleSerTyrlleuAsnArgHisAspAsnGlnAsnLeuArgVal	140
Qy	905	GCTAAAAATCTCCATCCCGGTTTAAACACACCCGGCGAGTTTGAGGATTTCTTCCCGGCGAGC	964
Db	141	AlaLysIleSerMetProValAsnThrProGlyGlnPheGluAspPhePheProAlaSer	160
Qy	965	AGCGGACCAATCATCTTTCGAGGGCTTCAGAGGAATACGTTGGAGGCGCGCTTC	1024
Db	161	SerArgAspGlnSerSerTyrlleuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPhe	180

QY 323 ACAGTGGCCGCAACCCGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnProGlyGluGlyGluAspGluGlnProArgProIle 104
QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAAGGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCAAGCTGGACCGGAGGAGCGGTGAAGAGAAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
QY 431 GACTGGAGACAACCAAGAGAAGAT-----TGGAGGCGCAACAGTCATCAGCAGCCACGG 484
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
QY 485 AAAATAAGGCCGAAGAGAGAGAAACAAGAGTGGGGAAACACACGAGTAGCCATGTG 544
Db 165 ArgLysGlnGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGA-----GAAACATCTCGGAACAACCTTTCTACTCCCGTCAAGGCGGTTTAGC 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCCGCTACGGGAACCAAAACGGTAGATCCGGCTCTCGAGAGTTTGACCAAGGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrglyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACAT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrglyIleGluGluPheAsnSerLysProAsnThr 244
QY 719 CTTGTCTTCCCAAGCAGCTGATGTGATAACATCTTGTATTTCAGCAAGGCGCAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrglyLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTGACCGTAGCAATCAATACAGAAAGAGCTTTAATCTTCAGGAGGCCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrglyArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGAAATCCCATCCGGTTTCATTTCTACATCTGAACCGCCATGACCAACAGAACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrglyValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGCTAAATCTCCATGCCGTTAAACACACCGCGCAGTTTGAGATTCTTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
QY 959 GCGAGCAGCGGACCAATCATCTACTTCAGGGCTTCAGCAGGAATAGTTGGAGGCC 1018
Db 325 SerSerThrGluAlaGlnGlnSerTyrglyLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGCTGCTGTAGAGAATGCGAGGAGT 1078
Db 345 SerTyrglyAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGly 363
QY 1079 GAGCAAGAGAGAGAGCGGACGCGGATGAGTACTCGAGTAGTCAGAACATGAAGGA 1138
Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGGTCAAGAGGACGAGTTGAAGAACTTACTAAGCAGCGTAAATCCGTC 1198
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY 1199 TCAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCCATCACTTGTAGAGAAGGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTCTAACTTTTGGGAAGTTATTGAGGTGAAGCCAGACAGAAAGAAC 1318
Db 412 AspProIleTyrglySerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGATCAAGAGGAGCTTTG 1378

Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAAGAACTGGA 1438
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAATCTGCTGCTGTAAAGAAAGAGCAACACAGAGGGGCGCGGGAAGAG 1498
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAGGAGGGAAGTAACACAGAGAGTGTAGGTACACAGCGAGG 1558
Db 490 Pro-----LeuGluValArgLysTyrglyArgAlaGlu 499
QY 1559 TTGAAGGAAGGCGATGTGTTTCATTCATGCGCAGAGCTCATCCAGTAGCCATCAACGTTCC 1618
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrglyProValValAlaAsnAlaThr 519
QY 1619 TCCGAATCCATCTGCTGCTGCTGCTATCAACGCTGAAACCAACACAGAGATCTTCTCTT 1678
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGCAATGTGTATAGCACCATGATAGAGACGACGAGCGAAGGATTTAGCATTC 1738
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CCTGGGTCGGGTGAACAAGTTGAGAAAGCTCATCAAAAAACAGAAAGAAATCTCATTGTG 1798
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrglyPheVal 579
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAG 1858
Db 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTTCCACTCCTTTCAATTTTGAAG 1918
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
Db 603 AlaPhe 604

Search completed: August 24, 2005, 10:52:48
Job time : 167.012 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 2122.81 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcacatactagtagcccc.....taaaagaatcatgtttttgtt 717

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	717	100.0	717	6	AR257468 Sequence
2	717	100.0	717	8	L77197 Arachis hyp
3	713.8	99.6	743	6	AR257462 Sequence
4	705	98.3	810	8	AY722689 Arachis h
5	644.6	89.9	842	8	AY581853 Arachis h
6	598.8	83.5	678	8	AY581854 Arachis h
7	474	66.1	474	6	BD107899 Methods a
8	472.4	65.9	474	6	AX155333 Sequence
9	470.8	65.7	474	6	BD172108 Peanut al
10	462	64.4	472	8	AY007229 Arachis h
11	462	64.4	682	6	AX148740 Sequence
12	462	64.4	1162	6	AX148738 Sequence
13	419.4	58.5	519	8	AY158467 Arachis h
14	415.2	57.9	510	8	AY117434 Arachis h
15	404.8	56.5	502	8	AF366560 Arachis h
16	220.8	30.8	634	8	AF366561 Arachis h
17	202.2	28.2	742	8	AY722690 Arachis h
18	197.8	27.6	627	8	AF092846 Arachis h
19	114.8	16.0	712	8	AF091737 Arachis h

20	108	15.1	581	8	AY722691	Arachis h
21	80	11.2	80	6	AX148744	Sequence
22	73.2	10.2	684	8	LACONGLD	X53523 L. angustif
23	65.6	9.1	777	6	AR066463	Sequence
24	65.6	9.1	777	8	GMU71195	U71195 Glycine max
25	64	8.9	777	6	AR066464	Sequence
26	63.4	8.8	770	6	AR107077	Sequence
27	63.4	8.8	770	6	AR304633	Sequence
28	63.4	8.8	770	6	BD075791	BD075791 Lunacin p
29	63.4	8.8	770	6	AF005030	Glycine m
30	62	8.6	62	6	AX148745	Sequence
31	59.2	8.3	477	6	AX839974	Sequence
32	59.2	8.3	477	6	AX840499	Sequence
33	58.8	8.2	7218	6	I66494	Sequence 14
34	57.6	8.0	82139	3	AC115684	AC115684 Dictyoste
35	56.4	7.9	127640	8	AC146555	Medicago
36	55	7.7	160542	2	CR627487	CR627487 Danio rer
37	54.4	7.6	468	6	AX839972	Sequence
38	54.4	7.6	468	6	AX840497	Sequence
39	54.4	7.6	723	6	AR066462	Sequence
40	54.4	7.6	724	8	GMU71194	U71194 Glycine max
41	53.6	7.5	174311	2	CR678065	CR678065 Danio rer
42	52.8	7.4	168571	9	CNS01DUF	AL133249 BAC seque
43	52.8	7.4	197146	5	BX470199	BX470199 Zebrafish
44	52.8	7.4	230976	2	BX927201	BX927201 Danio rer
45	52.6	7.3	199551	2	AC006281	AC006281 Plasmodiu

ALIGNMENTS

RESULT 1	AR257468	Sequence 20 from patent US 6486311.	717 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR257468	Sequence 20 from patent US 6486311.	717 bp	DNA	linear	PAT 20-DEC-2002
DEFINITION	AR257468	Sequence 20 from patent US 6486311.	717 bp	DNA	linear	PAT 20-DEC-2002
ACCESSION	AR257468	Sequence 20 from patent US 6486311.	717 bp	DNA	linear	PAT 20-DEC-2002
VERSION	AR257468.1	GI:27307481	717 bp	DNA	linear	PAT 20-DEC-2002
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 717)					
AUTHORS	Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,					
	Helms,R.M. and Bannon,G.A.					
TITLE	Peanut allergens and methods					
JOURNAL	Patent: US 6486311-A 20 26-NOV-2002;					
FEATURES	Location/Qualifiers					
source	1..717					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN	Query Match	100.0%;	Score 717;	DB 6;	Length 717;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-171;		
	Matches 717;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCTACCATCTAGTAGTACGCTCGCTTTCCTCTGCTGCCACGCGATCTCGAGGCA	60		
DB	1	GCTACCATCTAGTAGTACGCTCGCTTTCCTCTGCTGCCACGCGATCTCGAGGCA	60		
QY	61	GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAAGCTGAG	120		
DB	61	GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAAGCTGAG	120		
QY	121	GCCTCGAGGCAACATCTCATGCAGAGATCCAACTGACGAGGATTCATATCAACGGGA	180		
DB	121	GCCTCGAGGCAACATCTCATGCAGAGATCCAACTGACGAGGATTCATATCAACGGGA	180		
QY	181	CCCGTACAGCCCTAGTCAGGATCCGTCACGCCCTAGTCATATGATCGGAGGCGCTGG	240		
DB	181	CCCGTACAGCCCTAGTCAGGATCCGTCACGCCCTAGTCATATGATCGGAGGCGCTGG	240		
QY	241	ATCCTCTAGCACCAGAGAGAGGTGTGCAATGAGCTGAAGAGTTTGTGAGAACCAACCAAG	300		

```
Db      241  |||||ATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACCAAG 300
Qy      301  GTGCTATGTGCGAGGCAATGCAACAGATCATGGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Db      301  GTGCTATGTGCGAGGCAATGCAACAGATCATGGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Qy      361  GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCAGGCTTAG 420
Db      361  GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCAGGCTTAG 420
Qy      421  GGCAACACAGCGTTGCGACTTGGAACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
Db      421  GGCAACACAGCGTTGCGACTTGGAACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
Qy      481  TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db      481  TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Qy      541  GTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600
Db      541  GTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Qy      601  ACTAAGCGAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Db      601  ACTAAGCGAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Qy      661  TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATATAAAAGATCATGTTTTGTT 717
Db      661  TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATATAAAAGATCATGTTTTGTT 717

RESULT 2
ARQALLII
LOCUS   Arachis hypogaea (clone Ara h II p38) allergen II gene, polyA
DEFINITION
L77197.1 GI:1236995
VERSION 177197.1
KEYWORDS allergen; conglutinin; seed storage protein.
SOURCE  Arachis hypogaea (peanut)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 717)
Stanley,J.S.
The major peanut allergen Ara h II is a seed storage protein with
multiple IgE-binding epitopes
Unpublished (1996)
JOURNAL
COMMENT Original source text: Arachis hypogaea (strain Florunner) (clone:
Ara h II p38) DNA.
FEATURES
source Location/Qualifiers
polyA_signal 1..717
/mol_type="genomic DNA"
polyA_signal 562..567
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 717; DB 8; Length 717;
Best Local Similarity 100.0%; Pred.No.1.7e-171; Indels 0; Gaps 0;
Matches 717; Conservative 0; Mismatches 0;

Qy      1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTGCGAGGCA 60
Db      1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTGCGAGGCA 60
Qy      61 GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
```

```
Db      61  GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
Qy      121  GGCCTGCGAGCAACATCTCATGCAAGATCAACGCTGACGAGGATTCATATGAACGGGA 180
Db      121  GGCCTGCGAGCAACATCTCATGCAAGATCAACGCTGACGAGGATTCATATGAACGGGA 180
Qy      181  CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGG 240
Db      181  CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGG 240
Qy      241  ATCCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACCAAG 300
Db      241  ATCCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACCAAG 300
Qy      301  GTGCAATGTCGAGGCAATGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Db      301  GTGCAATGTCGAGGCAATGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Qy      361  GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCAGGCTTAG 420
Db      361  GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCAGGCTTAG 420
Qy      421  GGCACCAACAGCGTTGGGACTTGGAACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
Db      421  GGCACCAACAGCGTTGGGACTTGGAACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
Qy      481  TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db      481  TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Qy      541  GTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Db      541  GTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Qy      601  ACTAAGCGAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Db      601  ACTAAGCGAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Qy      661  TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATATAAAAGATCATGTTTTGTT 717
Db      661  TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATATAAAAGATCATGTTTTGTT 717

RESULT 3
ARQALLII
LOCUS   Arachis hypogaea (clone Ara h II p38) allergen II gene, polyA
DEFINITION
L77197.1 GI:1236995
VERSION 177197.1
KEYWORDS allergen; conglutinin; seed storage protein.
SOURCE  Arachis hypogaea (peanut)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 743)
Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannon,G.A.
Peanut allergens and methods
Patent: US 6486311-A 1 26-NOV-2002;
FEATURES
source Location/Qualifiers
polyA_signal 1..743
/mol_type="genomic DNA"

ORIGIN
Query Match 99.6%; Score 713.8; DB 6; Length 743;
Best Local Similarity 99.7%; Pred.No.1.1e-170; Indels 0; Gaps 0;
Matches 715; Conservative 0; Mismatches 2;

Qy      1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTGCGAGGCA 60
Db      6 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTGCGAGGCA 65
Qy      61 GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
```

```
Db 66 GCATGGGAATCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAG 125
Qy 121 GCCCTGCGAGCAACATCTCATGCGAGAAGATCCAAACGTCGACGAGGATTCATATGAACGGGA 180
Db 126 GCCCTGCGAGCAACATCTCATGCGAGAAGATCCAAACGTCGACGAGGATTCATATGAACGGGA 185
Qy 181 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 186 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 245
Qy 241 ATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGATTTGAGAACACCAAG 300
Db 246 ATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGATTTGAGAACACCAAG 305
Qy 301 GTGCATGTGCGAGCATTTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 360
Db 306 GTGCATGTGCGAGCATTTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 365
Qy 361 GCACAGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTGGCGCTTAG 420
Db 366 GCACAGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTGGCGCTTAG 425
Qy 421 GGACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
Db 426 GGACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 485
Qy 481 TATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATCTATG 540
Db 486 TATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATCTATG 545
Qy 541 GTTATGTTAGTTTGTGTAATAATAAAGATCATCACTATATGAATGTTGATCGTTTA 600
Db 546 GTTATGTTAGTTTGTGTAATAATAAAGATCATCACTATATGAATGTTGATCGTTTA 605
Qy 601 ACTAAGCAAGCTTAGTTTATATGACACCTTTAGAGTGTCTTTATGCGGTGTTCTATGT 660
Db 606 ACTAAGCAAGCTTAGTTTATATGACACCTTTAGAGTGTCTTTATGCGGTGTTCTATGT 665
Qy 661 TTTGTTGCTCAGAGTTGTAAACCATCTTGAAATAATAATAAAGATCATGTTTTGTT 717
Db 666 TTTGTTGCTCAGAGTTGTAAACCATCTTGAAATAATAATAAAGATCATGTTTTGTT 722
```

```
RESULT 4
AY722689
LOCUS Arachis hypogaea 2S protein 1 mRNA linear PLN 19-SEP-2004
DEFINITION Arachis hypogaea 2S protein 1 mRNA, partial cds.
ACCESSION AY722689
VERSION AY722689.1 GI:52001226
KEYWORDS
SOURCE
ORGANISM
```

```
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 810)
Yan, Y., Wang, L. and Huang, S.
cDNA clone of peanut seed storage protein gene
Unpublished
```

```
2 (bases 1 to 810)
Yan, Y., Wang, L. and Huang, S.
Direct Submission
Submitted (30-JUL-2004) Department of Biology and Biotechnology,
Sun Yat-sen University, Guangzhou, Guangdong 510275, China
LOCATION/Qualifiers
```

```
FEATURES
source
1..810
/organism="Arachis hypogaea"
/mol_type="mRNA"
/db_xref="taxon:3818"
<1..542
/note="seed storage protein"
```

CDS

```
ORIGIN
Query Match 98.3%; Score 705; DB 8; Length 810;
Best Local Similarity 99.9%; Pred. No. 1.9e-168;
Matches 716; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
/codon_start=3
/product="2S protein 1"
/protein_id="AAU21494.1"
/db_xref="GI:52001227"
/translation="GRGKHTLFTNHTSIITTTMAKLTILVALFLAAAHASARQOM
ELQDRRCQSOLERANLRPCOEHLMOKIOEDSDYERDPYSPQDPYSPSPYDRRGAG
SSQHQRCNCNELNEFENNRCMEALQOIMENQSDRLQGRQOQPKRELRLNPQQCG
LRAPQRCDLQDVGSGRDY"
```

```
Qy 1 GCTCACATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCCCACGATCTGCGAGGCA 60
Db 68 GCTCACATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCCCACGATCTGCGAGGCA 127
Qy 61 GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
Db 128 GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 187
Qy 121 GCCTCTGCGAGCAACATCTCATGCGAGAAGATCCAAACGTCGAGAGGATTCATATGAACGGGA 180
Db 188 GCCTCTGCGAGCAACATCTCATGCGAGAAGATCCAAACGTCGAGAGGATTCATATGAACGGGA 247
Qy 181 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 248 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 307
Qy 241 ATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGATTTGAGAACACCAAG 300
Db 308 ATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGATTTGAGAACACCAAG 367
Qy 301 GTGCATGTGCGAGCATTTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 360
Db 368 GTGCATGTGCGAGCATTTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 427
Qy 361 GCACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCCCTCAACAGTGGCGCTTAG 420
Db 428 GCACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCCCTCAACAGTGGCGCTTAG 487
Qy 421 GGACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACAGATCTAAACACC 480
Db 488 GGACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACAGATCTAAACACC 547
Qy 481 TATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATCTATG 540
Db 548 TATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATCTATG 607
Qy 541 GTTATGTTAGTTTGTGTAATAATAAAGATCATCACTATATGAATGTTGATCGTTTA 600
Db 608 GTTATG-TTATGTTTGTGTAATAATAAAGATCATCACTATATGAATGTTGATCGTTTA 666
Qy 601 ACTAAGCAAGCTTAGTTTATATGACACCTTTAGAGTGTCTTTATGCGGTGTTCTATGT 660
Db 667 ACTAAGCAAGCTTAGTTTATATGACACCTTTAGAGTGTCTTTATGCGGTGTTCTATGT 726
Qy 661 TTTGTTGCTCAGAGTTGTAAACCATCTTGAAATAATAATAAAGATCATGTTTTGTT 717
Db 727 TTTGTTGCTCAGAGTTGTAAACCATCTTGAAATAATAATAAAGATCATGTTTTGTT 783
```

```
RESULT 5
AY581853
LOCUS Arachis hypogaea seed storage protein SSP1 mRNA linear PLN 01-MAY-2004
DEFINITION Arachis hypogaea seed storage protein SSP1 mRNA, partial cds.
ACCESSION AY581853
VERSION AY581853.1 GI:46560479
KEYWORDS
SOURCE
ORGANISM
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 842)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
cDNA cloning of peanut seed storage protein
Unpublished
REFERENCE
1 (bases 1 to 842)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
Direct Submission
Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
Guangdong 510275, China
Location/Qualifiers
1..842
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="SanYou523"
/db_xref="taxon:3818"
<1..-564
/codon_start=1
/product="seed storage protein SSP1"
/protein_id="AA00598.1"
/db_xref="GI:46560480"
/translation="DHTLFNTHSIITTTTMAKLITLVALFLAAHASARQWELQG
DRCQSLERANLRPCEQHLMOKIQORDEDSYGRDPSPQDPYSPQDPDRDPYSPS
PYDRRGAGSGHOCRCNELNEFNORCMCEALQIMENQSDRLQGRQEQFKREL
RNLPQQGLRAPORCLEVESGGRDY"
ORIGIN
Query Match 89.9%; Score 644.6; DB 8; Length 842;
Best Local Similarity 93.9%; Pred. No. 4.5e-153;
Matches 707; Conservative 0; Mismatches 9; Indels 37; Gaps 2;
QY 1 GCTCACATACTAGTAGCCCTGCCCTTTTCCTCGTGCCTGCCACGCACTCTCGAGGCA 60
DB 54 GCTCACATACTAGTAGCCCTGCCCTTTTCCTCGTGCCTGCCACGCACTCTCGAGGCA 113
QY 61 GCAGTGGGAACCTCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
DB 114 GCAGTGGGAACCTCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTAG 173
QY 121 GCCTCGGAGCAACATCTCATGCAGAAAGATCCAACTGAGGAGATTCATATGAACGGGA 180
DB 174 GCCTCGGAGCAACATCTCATGCAGAAAGATCCAACTGAGGAGATTCATATGAACGGGA 233
QY 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCGCTGG 240
DB 234 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCGCTGG 293

QY 565 AAAGATCATCATATATGAATGTTGATCGTGTTAACTAAGCAAGCTTAGGTTATATG 624
DB 653 AAAGATCATCATATATGAATGTTGATCGTGTTAACTAAGCAAGCTTAGGTTATATG 712
QY 625 AGCACCTTTAGAGTGCTTTTATGGCGTGTCTATGTTTGTCTGCTGACAGTTGTAACCA 684
DB 713 AGCACCTTTAGAGTGCTTTTATGGCGTGTCTATGTTTGTCTGCTGACAGTTGTAACCA 772
QY 685 TCTTGAATAATAATAAAAGATCATGTTTGT 717
DB 773 TCTTGAATAATAATAAAAGATCATGTTTGT 805
RESULT 6
AY581854 678 bp mRNA linear PLN 01-MAY-2004
LOCUS Arachis hypogaea seed storage protein SSP2 mRNA, partial cds.
DEFINITION AY581854
ACCESSION AY581854
VERSION AY581854.1 GI:46560481
KEYWORDS Arachis hypogaea (peanut)
SOURCE Arachis hypogaea
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 678)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
cDNA cloning of peanut seed storage protein
Unpublished
REFERENCE
2 (bases 1 to 678)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
Direct Submission
Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
Guangdong 510275, China
Location/Qualifiers
1..678
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="SanYou523"
/db_xref="taxon:3818"
<1..-528
/codon_start=1
/product="seed storage protein SSP2"
/protein_id="AA00599.1"
/db_xref="GI:46560482"
/translation="DHTLFNTHSIITTTTMAKLITLVALFLAAHASARQWELQ
DRCQSLERANLRPCEQHLMOKIQORDEDSYGRDPSPQDPYSPQDPDRRGAGSGH
OERCCNELNEFNORCMCEALQIMENQSDRLQGRQEQFKRELRLNPQQGLRAP
QRCDDLVSSEGGGRDY"
ORIGIN
Query Match 83.5%; Score 598.8; DB 8; Length 678;
Best Local Similarity 99.4%; Pred. No. 2e-141;
Matches 622; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1 GCTCACATACTAGTAGCCCTGCCCTTTTCCTCGTGCCTGCCACGCACTCTCGAGGCA 60
DB 54 GCTCACATACTAGTAGCCCTGCCCTTTTCCTCGTGCCTGCCACGCACTCTCGAGGCA 113
QY 61 GCAGTGGGAACCTCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
DB 114 GCAGTGGGAACCTCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTAG 173
QY 121 GCCTCGGAGCAACATCTCATGCAGAAAGATCCAACTGAGGAGATTCATATGAACGGGA 180
DB 174 GCCTCGGAGCAACATCTCATGCAGAAAGATCCAACTGAGGAGATTCATATGAACGGGA 233
QY 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCGCTGG 240
DB 234 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCGCTGG 293

241 ATCTCTCAGACCAAGAGAGGTTGCAATGAGCTGAAAGAGTTTGAGAACCAAAAG 300
294 ATCTCTCAGACCAAGAGAGGTTGCAATGAGCTGAAAGAGTTTGAGAACCAAAAG 353
301 GTGCATGTGCGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGGAG 360
354 GTGCATGTGCGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGGAG 413
361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGGCTTTAG 420
414 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGGCTTTAG 473
421 GGACACACAGGTTGCGACCTTGGACCTCGAAGTGGCGGCGAGACAGATCTAAACACC 480
474 GGACACACAGGTTGCGACCTTGGACCTCGAAGTGGCGGCGAGACAGATCTAAACACC 533
481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAATAG-CTTATATATAAGCTATTATCTAT 539
534 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAATATCTTATATAAGCTATTATCTAT 593
540 GGTATGTTAGTTTTCGTTAATAAAGATCATCACTATATGAATGTGTGATCGGTT 599
594 GGTATG-TTAGTTTTCGTTAATAAAGATCATCACTATATGAATGTGTGATCGGTT 652
600 AACTAAGCAAGCTTAGTTTATGA 625
653 AACTAAGCAAGCTTAGTTTATGA 678

RESULT 7
BD107899
LOCUS BD107899 474 bp DNA linear PAT 18-SEP-2002
DEFINITION Methods and reagents for decreasing allergic reactions.
ACCESSION BD107899
VERSION BD107899.1 GI:23202717
KEYWORDS JP 2002501748-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 474)
Sosin, H., Banon, G.A., Jr, W.A.B. and Samphson, H.A.
Methods and reagents for decreasing allergic reactions
Patent: JP 2003501748-A 2 22-JAN-2002;
UNIVERSITY OF ARKANSAS, MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY
UNIVERSITY OF NEW YORK, HOWARD SOSIN
OS Arachis L. (Peanut)
PN JP 2002501748-A/2
PD 22-JAN-2002
PF 29-JAN-1999 JP 2000529437
PR 31-JAN-1998 US 60/073283,13-FEB-1998 US 60/074590 PR
13-FEB-1998 US 60/074624,13-FEB-1998 US 60/074633 PR
27-AUG-1998 US 09/141220
PI HAWADO SOSIN, GARY A. BANON, WESLEY A. BIRX JR, HYU A. SAMPHON PC
C12N15/09, A01H5/00, A01K67/027, A61K39/35, C07K14/37, C07K14/415, PC
C07K14/435,
PC C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/00, C12N5/00, C12N5/
00
CC Methods and reagents for decreasing allergic reactions FH
Key
FT source 1..474
FT Location/Qualifiers
1..474 /organism='Arachis L. (Peanut)'.
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

FEATURES
source
Query Match 66.1%; Score 474; DB 6; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e-109;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

2 CTCACCATACTAGTAGCCCTTCGCCCTTTCTCTCGTCCGCCACGCACTCTCGAGGCAG 61
1 CTCACCATACTAGTAGCCCTTCGCCCTTTCTCTCGTCCGCCACGCACTCTCGAGGCAG 60
62 CAGTGGGAACCTCCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 121
61 CAGTGGGAACCTCCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 120
122 CCTCGAGCAACATCTCATGCAAGAGATCCAAAGTGAAGAGGATTCATATGAACGGGAC 181
121 CCTCGAGCAACATCTCATGCAAGAGATCCAAAGTGAAGAGGATTCATATGAACGGGAC 180
182 CCTCGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGA 241
181 CCTCGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGA 240
242 TCCTCTCAGCAACCAAGAGAGGTTTGCATAGCTGAACGAGTTTGAGAAACCAACAGG 301
241 TCCTCTCAGCAACCAAGAGAGGTTTGCATAGCTGAACGAGTTTGAGAAACCAACAGG 300
302 TGCATGTGAGGAGCATTCGCAACAGATCATGAGAGAACAGAGCGATAGTTGCGAGGGGAG 361
301 TGCATGTGAGGAGCATTCGCAACAGATCATGAGAGAACAGAGCGATAGTTGCGAGGGGAG 360
362 CAACAGGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTGCAGGCTTAGG 421
361 CAACAGGCAACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTGCAGGCTTAGG 420
422 GCACCAACAGGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGATACTAA 475
421 GCACCAACAGGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGATACTAA 474

RESULT 8
AX155333
LOCUS AX155333 474 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6 from Patent WO0140264.
ACCESSION AX155333
VERSION AX155333.1 GI:14536768
KEYWORDS Arachis hypogaea (peanut)
SOURCE ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1
REFERENCE
AUTHORS Bannon, G.A., Burks, W.A., Caplan, M.J., Sampson, H. and Sosin, H.
TITLE Peptide antigens
JOURNAL Patent: WO 0140264-A 6 07-JUN-2001;
Panacea Pharmaceuticals, LLC (US); The University of Arkansas (US)
(US); MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK
LOCATION/Qualifiers
1..474 /organism='Arachis hypogaea'
/mol_type='unassigned DNA'
/db_xref='taxon:3818'

FEATURES
source
Query Match 65.9%; Score 472.4; DB 6; Length 474;
Best Local Similarity 99.8%; Pred. No. 2.9e-109;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 60
Db 11 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 70

Qy 61 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCCAGAGGGCGAACCTGAG 120
Db 71 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCCAGAGGGCGAACCTGAG 130

Qy 121 GCCCTGCGAGCAACATCTCATGTCAGAGAGATCCAAAGCTGACGAGATTCATATGAACGGGA 180
Db 131 GCCCTGCGAGCAACATCTCATGTCAGAGAGATCCAAAGCTGACGAGATTCATATGAACGGGA 190

Qy 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 240
Db 191 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 250

Qy 241 ATCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGAGTTTCAGAACACCAAG 300
Db 251 ATCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGAGTTTCAGAACACCAAG 310

Qy 301 GTGCATGTGCGAGGATTCGAAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 360
Db 311 GTGCATGTGCGAGGATTCGAAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 370

Qy 361 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTTCAACAGTGGCGCTTAG 420
Db 371 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTTCAACAGTGGCGCTTAG 430

Qy 421 GGCAACACAGCTTGGCAGCTTGGAGCTGCAAGTGGCGGCGAG 462
Db 431 GGCAACACAGCTTGGCAGCTTGGAGCTGCAAGTGGCGGCGAG 472

RESULT 11
AXI48740
LOCUS AXI48740 582 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 3 from Patent WO0136621.
ACCESSION AXI48740
VERSION AXI48740.1 GI:14347294
KEYWORDS Arachis hypogaea (peanut)
SOURCE Arachis hypogaea
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE 1
Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
Down-regulation and silencing of allergen genes in transgenic
peanut seeds
Patent: WO 0136621-A 3 25-MAY-2001;
Alabama A & M University (US)
JOURNAL Location/Qualifiers
FEATURES
source
1. .682
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"

CDS
110. .733
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41202.1"
/db_xref="GI:14347293"
/translation="MAKLTILVALALFLAAHASARQWELQDRRCQSLERANLRP
CEQHLMKIQRDEDSYERDFSPQDFSPSYDRRGAGSSQHQERCCNELNFEFENQ
RCMCEALQIMENQSDRLQROEQOFKELRNLPQCGLRAPQRCDLQVSGRRPR
IPPLITGSRRRHQSPYGNRRYSAMCLLPRAADGDGWPSPVAVDCSG"

ORIGIN
Query Match 54.4%; Score 462; DB 6; Length 682;
Best Local Similarity 100.0%; Pred. No. 1.3e-106;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 60
Db 67 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 126

Qy 61 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
Db 127 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 186

Qy 121 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGCTGACGAGATTCATATGAACGGGA 180
Db 187 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGCTGACGAGATTCATATGAACGGGA 246

Qy 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 240
Db 247 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 306

Qy 241 ATCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 307 ATCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 366

Qy 301 GTGCATGTGCGAGGATTCGAAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 360
Db 367 GTGCATGTGCGAGGATTCGAAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 426

Qy 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTTCAACAGTGGCGCTTAG 420
Db 427 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTTCAACAGTGGCGCTTAG 486

Qy 421 GGCAACACAGCTTGGCAGCTTGGAGCTGCAAGTGGCGGCGAG 462
Db 487 GGCAACACAGCTTGGCAGCTTGGAGCTGCAAGTGGCGGCGAG 528

RESULT 12
AXI48738
LOCUS AXI48738 1162 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 1 from Patent WO0136621.
ACCESSION AXI48738
VERSION AXI48738.1 GI:14347292
KEYWORDS Arachis hypogaea (peanut)
SOURCE Arachis hypogaea
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE 1
Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
Down-regulation and silencing of allergen genes in transgenic
peanut seeds
Patent: WO 0136621-A 1 25-MAY-2001;
Alabama A & M University (US)
JOURNAL Location/Qualifiers
FEATURES
source
1. .1162
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"
110. .733
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41202.1"
/db_xref="GI:14347293"
/translation="MAKLTILVALALFLAAHASARQWELQDRRCQSLERANLRP
CEQHLMKIQRDEDSYERDFSPQDFSPSYDRRGAGSSQHQERCCNELNFEFENQ
RCMCEALQIMENQSDRLQROEQOFKELRNLPQCGLRAPQRCDLQVSGRRPR
IPPLITGSRRRHQSPYGNRRYSAMCLLPRAADGDGWPSPVAVDCSG"

ORIGIN
Query Match 54.4%; Score 462; DB 6; Length 1162;
Best Local Similarity 100.0%; Pred. No. 1.4e-106;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 60
Db 118 GCTCACCATACTAGTACGCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGAGGCA 177

Qy 61 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
Db 178 GCAGTGGGAACCTCCAAGGAGACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 237

Qy 121 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGCTGACGAGATTCATATGAACGGGA 180

```
Db      238  GCCTCGCAGCAACATCTCATGCAGAGATCCCAACGTGACGAGATTCAATGAAACGGGA 297
Qy      181  CCGGTACAGCCCTAGTCAGATCCGTACACCTAGTCCATATGATCGGAGGCGCTGG 240
Db      298  CCGGTACAGCCCTAGTCAGATCCGTACACCTAGTCCATATGATCGGAGGCGCTGG 357
Qy      241  ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACAAAG 300
Db      358  ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACAAAG 417
Qy      301  GTGATGTGCGAGGCATGCAACAGATCATGTGAGAACACAGAGCGATAGTGTGAGGGGAG 360
Db      418  GTGCATGTGCGAGGCATGCAACAGATCATGTGAGAACACAGAGCGATAGTGTGAGGGGAG 477
Qy      361  GCACACGAGCACAGTTCAAGAGGAGCTCAGAACTTGCCCTCAACAGTGGCGCCTTAG 420
Db      478  GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCCTCAACAGTGGCGCCTTAG 537
Qy      421  GGCAACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 462
Db      538  GGCACACACAGGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 579

RESULT 13
LOCUS   AY158467               519 bp mRNA linear PLN 08-DEC-2002
DEFINITION Arachis hypogaea allergen Ara h 2.02 mRNA, complete cds.
ACCESSION AY158467
VERSION   AY158467.1 GI:26245446
KEYWORDS
SOURCE   Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
          Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 519)
          Chatel,J.-M., Bernard,H. and Orson,F.M.
          cDNA of peanut allergen Ara h2 isoform
          Unpublished
REFERENCE 2 (bases 1 to 519)
          Chatel,J.-M., Bernard,H. and Orson,F.M.
          Direct Submission
          Submitted (04-OCT-2002) LIAA, INRA, CEA, CE Saclay, DRM-SPI, Gif
          sur Yvette 91191, France
FEATURES             Location/Qualifiers
     source           1..519
                     /organism="Arachis hypogaea"
                     /mol_type="mRNA"
                     /db_xref="taxon:3818"
     CDS              1..519
                     /note="Ara h2 isoform"
                     /codon_start=1
                     /product="allergen Ara h 2.02"
                     /protein_id="AAN77576.1"
                     /db_xref="GI:26245447"
                     /translation="MAKLITLVALFLLAHASARQWELQDRRCQSLERANLRP
CEQHLQKIQDRDESYGRDPYSPQDPDRDPYSPQDPDRDPYSPYDRGAGSSQHQER
CCNELNEFNNQCMCEALQQIMENQSDRLQGRQQEQFKERLNLPQCCGLNAPQRC
DLEVESGGDRY"

ORIGIN
Query Match      58.5%; Score 419.4; DB 8; Length 519;
Best Local Similarity 91.8%; Pred. No. 9e-96;
Matches 469; Conservative 0; Mismatches 6; Indels 36; Gaps 1;

Qy      1 GCTCACCATACTAGTACCCCTCGCCCTTTTCCTCCTCGCTGCCACGCGCATCTCGGAGGCA 60
Db      9 GCTCACCATACTAGTACCCCTCGCCCTTTTCCTCCTCGCTGCCACGCGCATCTCGGAGGCA 68
Qy      61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGGCGACTCGAGAGGGCGAACTTGAG 120
```

```
Db      69  GCAGTGGGAATCTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACTTGAG 128
Qy      121  GCCTCGCAGCAACATCTCATGCAGAGATCCAAACGTGACGAGGATTCAATATGAACGGGA 180
Db      129  GCCTCGCAGCAACATCTCATGCAGAAATCCAAACGTGACGAGGATTCAATATGAGCGGA 188
Qy      181  CCCGTACAGCCCTAGTCAG-----GATCC 204
Db      189  CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCAGGACCCCGGACAGACGTGATCC 248
Qy      205  GTACAGCCCTAGTCATATGATCGAGAGGCGCTGATCCTCTCAGCACCAAGAGAGGTG 264
Db      249  GTACAGCCCTAGTCATATGATCGAGAGGCGCTGATCCTCTCAGCACCAAGAGAGGTG 308
Qy      265  TTGCAATGAGCTGAACAGGTTTGAGAAACAACAAAGGTGCATGTGCGAGGCAATTGCAACA 324
Db      309  TTGCAATGAGCTGAACAGGTTTGAGAAACAACAAAGGTGCATGTGCGAGGCAATTGCAACA 368
Qy      325  GATCATGGAGAACCCAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTCGAAGAG 384
Db      369  GATAATGGAGAACCCAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTCGAAGAG 428
Qy      385  GGAGCTCAGGAACCTTGCCTCAACAGTGGCGGCTTAGGGCACCAAGCGTTGCGACTTGA 444
Db      429  GGAGCTCAGGAACCTTGCCTCAACAGTGGCGGCTTAGGGCACCAAGCGTTGCGACTTGA 488
Qy      445  CGTCGAAAGTGGCGGCGAGACAGATATAA 475
Db      489  AGTCGAAAGTGGCGGCGAGACAGATATAA 519

RESULT 14
LOCUS   AY117434               510 bp mRNA linear PLN 01-JUN-2003
DEFINITION Arachis hypogaea allergen Ara h 2 isoform mRNA, partial cds.
ACCESSION AY117434
VERSION   AY117434.1 GI:31322016
KEYWORDS
SOURCE   Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
          Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 510)
          Becker,W.-M., Suhr,M., Lindner,B., Wicklein,D. and Lepp,U.
          Re-investigation of the major peanut allergen Arah2 on the
          molecular level
          Unpublished
REFERENCE 2 (bases 1 to 510)
          Becker,W.-M. and Suhr,M.
          Direct Submission
          Submitted (03-JUN-2002) B&M Allergology, Research Center Borstel,
          Parkallee 35, Borstel 23845, Germany
FEATURES             Location/Qualifiers
     source           1..510
                     /organism="Arachis hypogaea"
                     /mol_type="mRNA"
                     /db_xref="taxon:3818"
                     <1..510
                     /note="conglutin"
                     /codon_start=1
                     /product="allergen Ara h 2 isoform"
                     /protein_id="AAM78596.1"
                     /db_xref="GI:31322017"
                     /translation="LTLIVAPALFLLAHASARQWELQDRRCQSLERANLRPCEQ
HLMKIQDRDESYGRDPYSPQDPDRDPYSPYDRGAGSSQHQERCCN
ELNEFENQRCMCEALQQIMENQSDRLQGRQQEQFKERLNLPQCCGLRAPQRC
DLEVESGGDRY"

ORIGIN
Query Match      57.9%; Score 415.2; DB 8; Length 510;
Best Local Similarity 91.4%; Pred. No. 1e-94;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 268.476 Seconds

(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 717

Sequence: 1 gctcaccatactagtagccc.....taaaagatcatgtttgtt 717

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1980s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	100.0	717	2	AAT76615
2	717	100.0	717	2	AAT76614
3	474	66.1	474	2	AZ06383
4	474	66.1	474	4	AAS08538
5	474	66.1	474	8	ABX70606
6	474	66.1	474	10	ADG27517
7	462	64.4	682	4	AAF90337
8	462	64.4	1162	4	AAF90336
9	80	11.2	80	4	AAF90342
10	65.6	9.1	777	2	AAT86282
11	64	8.9	777	2	AAT86283
12	63.4	8.8	770	2	AAZ5346
13	62	8.6	62	4	AAF90343
14	59.2	8.3	477	10	ADH89268
15	59.2	8.3	477	12	ADG44003
16	54.4	7.6	468	10	ADH89266
17	54.4	7.6	468	12	ADG44001
18	54.4	7.6	723	2	AAT86281
19	51.4	7.2	5392	2	AAZ08720
20	49.8	6.9	62658	13	ABD33339

c 20 49.8 6.9 62658 13 ABD33339 Human can

c	21	49.6	6.9	52216	4	AAH28355	Aah28355 Nucleotid
c	22	49.6	6.9	52216	6	ABL50307	AbL50307 Human mus
c	23	49.6	6.9	169739	6	ABQ88186	Abq88186 Human ost
	24	49	6.8	129	4	AAF87134	Aaf87134 Lunasin c
	25	49	6.8	129	4	AAF83653	Aaf83653 Soybean l
	26	49	6.8	9407	4	AAF83654	Aaf83654 pPIC9K-lu
	27	49	6.8	9408	4	AAF87135	Aaf87135 Lunasin c
	28	48.2	6.7	7758	6	ABL33103	AbL33103 Human imm
	29	48	6.7	6509	6	ABL32226	AbL32226 Human imm
	30	47.4	6.6	97835	6	ABK84796	Abk84796 Human CDN
	31	45.6	6.4	2000	8	ADA71938	Ada71938 Rice gene
	32	45.6	6.4	6063	6	ABK28393	Abk28393 DNA trans
c	33	45.4	6.3	5823	6	ABK28382	Abk28382 DNA trans
	34	45.2	6.3	12174	6	ABQ88143	Abq88143 Human ost
	35	44.4	6.2	6050	10	ADB84181	AdB84181 Human lym
c	36	44.4	6.2	6385	4	AAS45353	AAk45353 Chemicall
c	37	44.4	6.2	6385	6	ABK28186	Abk28186 DNA trans
	38	44.4	6.2	110000	2	AAV21209_04	Continuation (5 of
	39	44	6.1	608	6	ABN62524	Abn62524 Human can
	40	43.8	6.1	7566	4	AAI07322	Aal07322 Human rep
	41	43.8	6.1	154068	11	ADP65423	Adp65423 Human seq
c	42	43.8	6.1	174448	11	ACN43946	Acn43946 Human gen
	43	43.6	6.1	60815	11	ACN43882	Acn43882 Human gen
	44	43.4	6.1	5919	6	ABL32389	AbL32389 Human imm
	45	43.4	6.1	5919	6	ABK31195	Abk31195 Signal tr

ALIGNMENTS

RESULT 1

AAAT76615

ID AAT76615 standard; cDNA to mRNA; 717 BP.

XX

AC AAT76615;

XX

DT 17-OCT-2003 (revised)

DT 29-DEC-1997 (first entry)

XX

DE Peanut allergen Ara hII cDNA clone.

XX

KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;

KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;

KW ELISA; analysis; Ara hII; ds.

XX

OS Arachis hypogaea; strain Florunner.

XX

FH Key Location/Qualifiers

FT CDS 2..475

FT polyA_signal 562..567

FT /*tag= a

FT /*tag= b

XX

PN WO9724139-A1.

XX

PD 10-JUL-1997.

XX

PF 23-SEP-1996; 96WO-US015222.

XX

PR 29-DEC-1995; 95US-0009455P.

XX

PR 04-MAR-1996; 96US-00610424.

XX

PA (UYAR-) UNIV ARKANSAS.

XX

PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

XX

DR WPI; 1997-363453/33.

XX

DR P-PSDB; AAN24153.

XX

PT Peanut allergens Ara hI and Ara hII - used for vaccination and in two-

XX site monoclonal antibody based ELISA.

PS Claim 31; Page 196; 354pp; English.

XX	This cDNA clone codes for the major peanut allergen Ara hII (AAW22153), which contains multiple IgE-binding epitopes. It was amplified from a peanut seed cDNA library using a primer (see AAT59683) based on an isolated Ara hII peptide (see AAW24151). The sequence shows significant homology with the conglutin family of seed storage proteins of other legumes. The gene is capable of producing a protein product in prokaryotic cells that is recognised by serum IgE from a large proportion of individuals with peanut hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used to raise monoclonal antibodies which are used in a specific two- site MAb ELISA for the detection of Ara hI or Ara hII (claimed). IgE-binding Ara hII antigen epitopes (see AAW24188- 93) may be used in vaccines to protect against allergic reactions to peanut allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to standardise OS field)									
XX	Sequence 717 BP; 217 A; 152 C; 184 G; 164 T; 0 U; 0 Other;									
SQ	Query Match 100.0%; Score 717; DB 2; Length 717; Best Local Similarity 100.0%; Pred. No. 5.9e-178; Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCACGCACTCGGAGGCA	60							
Db	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCACGCACTCGGAGGCA	60							
Qy	61	GCAGTGGGAATCCNAGGAGACAGAGAATGCCAGAGCAGCTCGAGAGGGCGAACCTGAG	120							
Db	61	GCAGTGGGAATCCNAGGAGACAGAGAATGCCAGAGCAGCTCGAGAGGGCGAACCTGAG	120							
Qy	121	GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACAGGATTCATATGAACGGGA	180							
Db	121	GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACAGGATTCATATGAACGGGA	180							
Qy	181	CCCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG	240							
Db	181	CCCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG	240							
Qy	241	ATCCTCTCAGCACCAAGAGAGGTGTGCNATGAGCTGAACAGATTGAGACACACCAAG	300							
Db	241	ATCCTCTCAGCACCAAGAGAGGTGTGCNATGAGCTGAACAGATTGAGACACACCAAG	300							
Qy	301	GTGCATGTGCGAGGCAATTGCAACAGATCATGCGAGAACCCAGAGCGATAGTTGCGAGGGAG	360							
Db	301	GTGCATGTGCGAGGCAATTGCAACAGATCATGCGAGAACCCAGAGCGATAGTTGCGAGGGAG	360							
Qy	361	GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCGCTCAAAGTCGGCGCTTAG	420							
Db	361	GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCGCTCAAAGTCGGCGCTTAG	420							
Qy	421	GGCACCAAGCGTTGCGACTTTGGAGCTGCGAAAGTGGCGGACAGACAGATACTAAACACC	480							
Db	421	GGCACCAAGCGTTGCGACTTTGGAGCTGCGAAAGTGGCGGACAGACAGATACTAAACACC	480							
Qy	481	TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	540							
Db	481	TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	540							
Qy	541	GTTATGTTTAGTTTTGGTAAATAAAGATCATCACTATATGAATGTTGATCGTGTTA	600							
Db	541	GTTATGTTTAGTTTTGGTAAATAAAGATCATCACTATATGAATGTTGATCGTGTTA	600							
Qy	601	ACTAAGGCAAGCTTAGGTTATATAGCACCTTTTAGAGTGCCTTTTATGGCGTTGTCTATGT	660							
Db	601	ACTAAGGCAAGCTTAGGTTATATAGCACCTTTTAGAGTGCCTTTTATGGCGTTGTCTATGT	660							
Qy	661	TTTGTTGCTGCAGAGTTGTAAACCATCTTGGAAATATATAAAAAAGATCATGTTTTGTT	717							
Db	661	TTTGTTGCTGCAGAGTTGTAAACCATCTTGGAAATATATAAAAAAGATCATGTTTTGTT	717							

AAAT76614	standard; cDNA to mRNA; 717 BP.
AAAT76614;	
17-OCT-2003	(revised)
29-DEC-1997	(first entry)
Peanut allergen Ara hII cDNA clone p38.	
Peanut; seed storage protein; allergen; allergy; hypersensitivity;	
vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;	
ELISA; analysis; Ara hII; ds.	
Arachis hypogaea; strain Florunner.	
Key	Location/Qualifiers
CDS	2..475
polyA_signal	/*tag= a
562..567	
/*tag= b	
W09724139-Al.	
10-JUL-1997.	
23-SEP-1996;	96WO-US015222.
29-DEC-1995;	95US-0009455P.
04-MAR-1996;	96US-00610424.
(UYAR-) UNIV ARKANSAS.	
Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;	
WPI; 1997-363453/33.	
Peanut allergens Ara hI and Ara hII - used for vaccination and in two-	
site monoclonal antibody based ELISA.	
Claim 31; Page 183-185; 354pp; English.	
This cDNA clone, designated p38, codes for the major peanut allergen Ara	
hII (AAW22164), which contains multiple IgE-binding epitopes. It was	
amplified from a peanut seed cDNA library using a primer (see AAT58683)	
based on an isolated Ara hII peptide (see AAW24151). The sequence shows	
significant homology with the conglutin family of seed storage proteins	
of other legumes. The gene is capable of producing a protein product in	
prokaryotic cells that is recognised by serum IgE from a large proportion	
of individuals with peanut hypersensitivity. Ara hII and Ara hI (see	
AAW24149-50) can be used to raise monoclonal antibodies which are used in	
a specific two-site MAb ELISA for the detection of Ara hI or Ara hII	
(claimed). IgE-binding Ara hII antigen epitopes (see AAW24188-93) may be	
used in vaccines to protect against allergic reactions to peanut	
allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to	
standardise OS field)	
Sequence 717 BP; 217 A; 152 C; 184 G; 164 T; 0 U; 0 Other;	

QY 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGG 240
 DB 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGG 240
 QY 241 ATCTCTCTCAGCACCACCAAGAGGCTGTGCAATGAGCTGAACGAGTTTCAGAACCAACCAAG 300
 DB 241 ATCTCTCTCAGCACCACCAAGAGGCTGTGCAATGAGCTGAACGAGTTTCAGAACCAACCAAG 300
 QY 301 GTGATGTGGAGGATTCGCAACAGATCATGAGAAACAGAGCGATAGGTTCAGAGGGAG 360
 DB 301 GTGATGTGGAGGATTCGCAACAGATCATGAGAAACAGAGCGATAGGTTCAGAGGGAG 360
 QY 361 GCACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
 DB 361 GCACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
 QY 421 GGACACACAGCGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
 DB 421 GGACACACAGCGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAAACACC 480
 QY 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
 DB 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
 QY 541 GTTATGTTTGTAGTTTGTGTAATAATAAAGATCATCACTATATGATGTGTCGTGTTA 600
 DB 541 GTTATGTTTGTAGTTTGTGTAATAATAAAGATCATCACTATATGATGTGTCGTGTTA 600
 QY 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTTGAAGTGGCTTTTATGCGGTGTCATGT 660
 DB 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTTGAAGTGGCTTTTATGCGGTGTCATGT 660
 QY 661 TTTGTTGCTGCAGAGTTGTTAAACCATCTTGAATAATAATAAAGATCATGTTTGT 717
 DB 661 TTTGTTGCTGCAGAGTTGTTAAACCATCTTGAATAATAATAAAGATCATGTTTGT 717

RESULT 3

AAZ06383
 ID AAZ06383 standard; DNA; 474 BP.
 AC AAZ06383;
 XX
 XX 17-OCT-2003 (revised)
 DT 09-NOV-1999 (first entry)
 XX
 XX Peanut allergen, Ara h 2.
 XX allergy; immune response; transgenic; allergen; epitope;
 KW immunoglobulin E; Ig E; binding site; peanut; ds.
 XX
 OS Arachis hypogaea.
 XX
 XX WO9398978-A1.
 XX
 PD 05-AUG-1999.
 XX
 XX 29-JAN-1999; 99WO-US002031.
 XX
 XX 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 27-AUG-1998; 98US-00141220.
 XX
 XX (UVAR-) UNIV ARKANSAS.
 PA (UNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) SOSIN H.
 XX
 XX Sosin H, Bannon GA, Burks AW, Sampson HA;
 XX WPI; 1999-479189/40.

DR P-PSDB; AAY15245.
 XX Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 XX
 XX Disclosure; Page 37; 46pp; English.
 XX
 CC This is the nucleotide sequence of the Ara h 2 protein from Arachis
 CC hypogaea. The Ara h 2 protein has 10 IgE (immunoglobulin E) binding
 CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
 CC AAY15276). By modifying the IgE binding sites the ability of the allergen
 CC to provoke an immune response is downregulated. The epitopes of the IgE
 CC binding sites can therefore be modified in genetically engineered plants
 CC and animals to elicit less of an allergic response. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;

Query Match 66.1%; Score 474; DB 2; Length 474;
 Best Local Similarity 100.0%; Pred. No. 3.3e-114;
 Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CTCAACATCTAGTAGCCCTCGCCCTTTTCTCTCTCGTCCGCCACGATCTCGAGGAG 61
 DB 1 CTCAACATCTAGTAGCCCTCGCCCTTTTCTCTCTCGTCCGCCACGATCTCGAGGAG 60
 QY 62 CAGTGGGAATCCAAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTGAGG 121
 DB 61 CAGTGGGAATCCAAAGAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTGAGG 120
 QY 122 CCTGTGGAGCAATCTCATGCAAGATCCACGTCGAGGAGATTCATATGAACGGGAC 181
 DB 121 CCTGTGGAGCAATCTCATGCAAGATCCACGTCGAGGAGATTCATATGAACGGGAC 180
 QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGAGCGCTGGA 241
 DB 181 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGAGCGCTGGA 240
 QY 242 TCCTCTCAGCACCACAGAGAGGTTTGCATAGCTGAACGAGTTTGAAGAACCAACAAAG 301
 DB 241 TCCTCTCAGCACCACAGAGAGGTTTGCATAGCTGAACGAGTTTGAAGAACCAACAAAG 300
 QY 302 TGCATGTGGAGCATTGCAACAGATCATGGAGAACCCAGAGCCATAGTTGCGAGGGGAGG 361
 DB 301 TGCATGTGGAGCATTGCAACAGATCATGGAGAACCCAGAGCCATAGTTGCGAGGGGAGG 360
 QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGCCCTTAGG 421
 DB 361 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGCCCTTAGG 420
 QY 422 GCACACAGCGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAA 475
 DB 421 GCACACAGCGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACTAA 474

RESULT 4

AA508538
 ID AA508538 standard; cDNA; 474 BP.
 AC AA508538;
 XX
 XX 23-OCT-2001 (first entry)
 DT
 XX DNA encoding anaphylactic antigen Ara h 2.
 DE
 XX Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse; ss.
 XX
 OS Mus sp.
 XX
 XX Key Location/Qualifiers
 FT 1..474
 FT CDS /*tag= a

```

FT  /product= "Peptide antigen Ara h 2"
XX WO200140264-A2.
XX
XX
XX
XX 07-JUN-2001.
XX
XX 06-DEC-2000; 2000WO-US033124.
XX
XX 06-DEC-1999; 99US-00455294.
XX 23-JUN-2000; 2000US-0213765P.
XX 27-SEP-2000; 2000US-0235797P.
XX
XX (PANA-) PANACEA PHARM LLC.
XX (UYAR-) UNIV ARKANSAS.
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
XX Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX
XX WPI; 2001-381378/40.
XX P-PSDB; AAU04707.
XX
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
XX the severity and/or number of allergic symptoms in individuals sensitive
XX to antigens, have reduced ability to bind immunoglobulin E.
XX
XX Disclosure; Fig 10; 100pp; English.
XX
XX The sequence represents the coding sequence of anaphylactic antigen Ara h
XX 2. Ara h 2 is an anaphylactic antigen (A), which was used to design
XX antigenic peptides having a reduced ability to bind immunoglobulin E
XX (Ige) as compared with the intact (A), or having a sequence substantially
XX identical to a portion of sequence of an antigen that includes at least
XX one Ige binding site, where at least one Ige binding site of the peptide
XX is altered. The antigenic peptides are used in a composition which is
XX useful for reducing risk or severity of allergic reaction to an antigen.
XX This is done by identifying an individual at risk of allergic reaction to
XX an antigen by identifying prior display of allergic symptoms when exposed
XX to the antigen, or a familial relationship with an individual who
XX previously displayed allergic symptoms when exposed to the antigen.
XX Following this an antigen-specific Ige present on one or more mast cells
XX or basophils in the individual's serum is identified. The individual is
XX then contacted with a peptide corresponding to a portion of the antigen,
XX which is selected, formulated, and delivered so that binding of the
XX peptide to antigen-specific Ige is reduced as compared with Ige binding
XX of intact antigen. The composition is also useful for treating and
XX preventing allergic reactions
XX
XX Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;
XX
XX Query Match 66.1%; Score 474; DB 4; Length 474;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-114;
XX Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCACTCGGAGCGAG 61
XX |
XX 1 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCACTCGGAGCGAG 60
XX |
XX 62 CAGTGGGAATCCAAAGGAGACAGAGATGCCAGAGCAGCTCCAGAGGGCGACCTGAGG 121
XX |
XX 61 CAGTGGGAATCCAAAGGAGACAGAGATGCCAGAGCAGCTCCAGAGGGCGAACTGAGG 120
XX |
XX 122 CCCTGCGAGCAACATCTCATGCGAGAGATCCAAACGTCGACGAGGATTCATATGAACGGGAC 181
XX |
XX 121 CCCTGCGAGCAACATCTCATGCGAGAGATCCAAACGTCGACGAGGATTCATATGAACGGGAC 180
XX |
XX 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCGTGA 241
XX |
XX 181 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCGTGA 240
XX |
XX 242 TCCTCTCAGCACCAAGAGAGGTTGTCATGAGCTGAACGAGTTTCAGACACCAACGAAG 301
XX |
XX 241 TCCTCTCAGCACCAAGAGAGGTTGTCATGAGCTGAACGAGTTTCAGACACCAACGAAG 300
XX |

```

```

QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGAGAACACGAGCGATAGGTTTGAGGGGAGG 361
DB |
DB 301 TGCATGTGCGAGGCATTGCAACAGATCATGAGAACACGAGCGATAGGTTTGAGGGGAGG 360
QY 362 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCCTTAGG 421
DB |
DB 361 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCCTTAGG 420
QY 422 GCACCACACGCTTGGGACTTGGACGTCGAAAGTGGCGGCAGAGACAGATACTAA 475
DB |
DB 421 GCACCACACGCTTGGGACTTGGACGTCGAAAGTGGCGGCAGAGACAGATACTAA 474

RESULT 5
ABX70606
ID ABX70606 standard; cDNA; 474 BP.
XX
AC ABX70606;
XX
DT 26-MAR-2003 (first entry)
XX
DE Peanut Ara h2 cDNA.
XX
KW Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; Ige binding site;
KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI; 2003-018765/01.
DR P-PSDB; ABU52464.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 10; Fig 40; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the Ige binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a cDNA
XX encoding a peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;
XX
XX Query Match 66.1%; Score 474; DB 8; Length 474;

```


Best Local Similarity 100.0%; Pred. No. 3.3e-114;		Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	2	CTCACCATACTAGTACGCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGACG 61	
Db	1	CTCACCATACTAGTACGCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGACG 60	
Qy	62	CAGTGGGAATCCAAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAATCTGAG 121	
Db	61	CAGTGGGAATCCAAAGAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAATCTGAG 120	
Qy	122	CCCTGCAGCAACATCTCATGCAGAGATCCAAAGTGCAGAGGATTCATATGAACGGAC 181	
Db	121	CCCTGCAGCAACATCTCATGCAGAGATCCAAAGTGCAGAGGATTCATATGAACGGAC 180	
Qy	182	CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGGA 241	
Db	181	CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGGA 240	
Qy	242	TCCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGATTTGAGAACCAACCAAGG 301	
Db	241	TCCTCTCAGCACCAGAGAGGTTGCAATGAGCTGAACGATTTGAGAACCAACCAAGG 300	
Qy	302	TGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCCAGAGCGATAGTTGCAGGGGAGG 361	
Db	301	TGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCCAGAGCGATAGTTGCAGGGGAGG 360	
Qy	362	CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGGCTCAACAGTGGCGCTTAGG 421	
Db	361	CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGGCTCAACAGTGGCGCTTAGG 420	
Qy	422	GCACCAGAGGTTGCGACTTGGAGCTGGAAGTGGCGGCGAGAGACAGATCTAA 475	
Db	421	GCACCAGAGGTTGCGACTTGGAGCTGGAAGTGGCGGCGAGAGACAGATCTAA 474	
RESULT 6			
ADG27517			
ID	ADG27517 standard; cDNA; 474 BP.		
XX	AC ADG27517;		
DT	26-FEB-2004 (first entry)		
XX	Peanut cDNA encoding Ara h2.		
DE	Peanut; plant; ss; gene; allergen; Ara h1; Ara h2; Ara h3;		
KW	Glycinin A2B1a; Jug nl; antiallergic; vulnerary;		
KW	anaphylactic food allergen; IgE; allergy; wound.		
XX	Arachis hypogaea.		
XX	US2003202980-A1.		
PN	30-OCT-2003.		
PD	18-MAR-2002; 2002US-00100303.		
XX	29-DEC-1995; 95US-0009455P.		
PR	23-SEP-1996; 96US-0071793P.		
PR	31-JAN-1998; 98US-0073283P.		
PR	13-FEB-1998; 98US-0074590P.		
PR	13-FEB-1998; 98US-0074633P.		
PR	29-JUN-1998; 98US-00106872.		
PR	27-AUG-1998; 98US-00141220.		
PR	13-NOV-1998; 98US-00191593.		
PR	29-JAN-1999; 99US-00240357.		
PR	11-FEB-1999; 99US-00241101.		
PR	11-FEB-1999; 99US-00248673.		
PR	02-MAR-1999; 99US-00248674.		
PR	02-MAR-1999; 99US-0122450P.		
PR	02-MAR-1999; 99US-0122452P.		
PR	02-MAR-1999;	99US-0122560P.	
PR	02-MAR-1999;	99US-0122565P.	
PR	02-MAR-1999;	99US-0122566P.	
PR	11-MAR-1999;	99US-00267719.	
PR	28-JAN-2000;	2000US-00494096.	
PR	16-MAR-2001;	2001US-0276822P.	
XX	(CAPL/) CAPLAN M J.		
PA	(SOSI/) SOSIN H B.		
PA	(SAMP/) SAMPSON H.		
PA	(BANN/) BANNON G A.		
PA	(BURK/) BURKS A W.		
PA	(COCK/) COCKRELL G.		
PA	(COMP/) COMPADRE C M.		
PA	(CONN/) CONNAUGHTON C.		
PA	(HELM/) HELM R M.		
PA	(KING/) KING N E.		
PA	(KOPP/) KOPPER R A.		
PA	(MALE/) MALEKI S J.		
PA	(RABJ/) RABJOHN P A.		
PA	(SHIN/) SHIN D S.		
PA	(STAN/) STANLEY J S.		
XX	Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;		
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
PI	Rabjohn PA, Shin DS, Stanley JS;		
XX	WPI; 2003-875632/81.		
DR	P-PSDB; ADG27518.		
XX	New modified anaphylactic food allergen comprising a cysteine residue		
PT	which has been modified so that it cannot participate in the disulfide		
PT	bond, useful for treating allergic reactions or wounds.		
XX	Example 10; SEQ ID NO 62; 194pp; English.		
PS	The invention relates to a modified anaphylactic food allergen whose		
CC	amino acid sequence is substantially identical to that of a natural		
CC	anaphylactic food allergen. The natural anaphylactic food allergen		
CC	includes at least one cysteine residue that participates in a disulphide		
CC	bond when the natural anaphylactic food allergen is in its native		
CC	conformation, except that the cysteine residue has been modified so that		
CC	it cannot participate in the disulphide bond. Also included are a method		
CC	of making a modified anaphylactic food allergen, a nucleotide molecule		
CC	encoding a modified anaphylactic food allergen defined above, a		
CC	nucleotide molecule for causing a site specific mutation in a gene		
CC	encoding a natural anaphylactic food allergen, a transgenic plant or		
CC	animal expressing a modified anaphylactic food allergen defined above, a		
CC	method of treating an individual by reducing the clinical response to a		
CC	natural anaphylactic food allergen by administering a modified		
CC	anaphylactic food allergen and an isolated fragment of peanut allergen		
CC	Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or		
CC	ADG27465. About 10-17% of the amino acids have been modified in at least		
CC	one IgE epitope or all the IgE epitopes recognised when the natural		
CC	anaphylactic food allergen is contacted with serum IgE from individual(s)		
CC	allergic to the natural anaphylactic food allergen. The invention		
CC	discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding		
CC	cDNAs), Soybean Glycinin A2B1a and IgE-binding epitopes of the English		
CC	walnut allergen Jug nl. The modified anaphylactic food allergen can be		
CC	used for treating allergic reactions or wounds. The present sequence		
CC	encodes a Peanut allergen of the invention.		
XX	Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;		
XX	Query Match 66.1%; Score 474; DB 10; Length 474;		
QQ	Best Local Similarity 100.0%; Pred. No. 3.3e-114;		
QQ	Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	2	CTCACCATACTAGTACGCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGACG 61	
Db	1	CTCACCATACTAGTACGCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGACG 60	
Qy	62	CAGTGGGAATCCAAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAATCTGAG 121	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 1681.67 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gtcacacatactagtagccc.....taaaagatcatgttttgtt 717

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hsc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_est8.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	686.4	95.7	788	6	CD038434
2	666	92.9	689	6	CD038213
3	555.2	77.4	732	6	CD038384
4	524	73.1	524	7	CD0897505
5	487.8	68.0	634	6	CD038323
6	315	43.9	405	6	CD038306
7	289	40.3	308	6	CD038459
8	283.4	39.5	375	7	CD0897504
9	223.8	31.2	696	6	CD038415
10	215.2	30.0	732	6	CD038394
11	208.8	29.1	661	6	CD038326
12	143.4	20.0	211	7	CD0897508
13	137.6	19.2	286	7	CD0897506
14	133.2	18.6	680	6	CD038092
15	131.8	18.4	351	7	CD0897507
16	129.6	18.1	714	6	CD038841
17	81	11.3	290	7	CD0897509
18	71.6	10.0	659	7	CN825259
19	67.6	9.4	305	6	CD038103
20	65.8	9.2	445	2	BF594963
21	63	8.8	719	2	BE661358
22	62.6	8.7	707	2	BE658140
23	62.4	8.7	479	4	BG237204
24	60.6	8.5	504	1	AJ498204

c

25	59.8	8.3	468	1	AI960869	8c91g03.Y
26	59.2	8.3	469	2	AW279501	AW279501 sf90d06.Y
27	59.2	8.3	469	2	AW318110	AW318110 sg60e10.Y
28	59.2	8.3	471	2	AW318155	AW318155 sg61b03.Y
29	59.2	8.3	474	2	AW318193	AW318193 sg62c02.Y
30	59.2	8.3	475	2	AW395515	AW395515 sg72b07.Y
31	59.2	8.3	475	2	AW395679	AW395679 sg73h12.Y
32	59.2	8.3	475	2	AW398039	AW398039 sg71d12.Y
33	59.2	8.3	477	1	AI759696	AI759696 sb62h10.Y
34	59.2	8.3	479	2	BF595027	BF595027 su74e01.Y
35	59.2	8.3	479	2	AW395742	AW395742 sg74g08.Y
36	59.2	8.3	481	2	AW396869	AW396869 sg64c10.Y
37	59.2	8.3	482	2	AW397141	AW397141 sg67d07.Y
38	59.2	8.3	486	1	AI941225	AI941225 sb66b06.Y
39	59.2	8.3	487	2	AW397497	AW397497 sg79f01.Y
40	59.2	8.3	489	2	BF596512	BF596512 su72c09.Y
41	59.2	8.3	490	1	AI748087	AI748087 sb48a09.Y
42	59.2	8.3	490	2	AW397370	AW397370 sg78a09.Y
43	59.2	8.3	490	2	AW398070	AW398070 sg71g12.Y
44	59.2	8.3	495	2	AW397013	AW397013 sg65h10.Y
45	59.2	8.3	521	2	AW395825	AW395825 sg75g05.Y

ALIGNMENTS

RESULT 1
CD038434
LOCUS
DEFINITION
CD038434
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD038434 788 bp mRNA linear EST 07-MAY-2003
UTPPI007 C04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI007_C04 5', mRNA sequence.

CD038434
EST.
CD038434.1 GI:30420272

Arachis hypogaea (peanut)

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

1 (bases 1 to 788)

Luo M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.

and Lynch, R.E.

Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

Contact: Baozhu Guo

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

Location/Qualifiers

1..788

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPPI007_C04"

/tissue_type="immature pods"

/dev_stage="R6"

/lab_host="XLI-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;

cDNA library was constructed from peanut cultivar A13;

(NCV1XAR4). A13 has resistance to Aspergillus infection

and drought tolerance. The immature pods that developed to

R6 stage were collected from different plants, and placed

into liquid N2 immediately and stored in -80oC freezer.

Total RNA was isolated with TRIzol-Reagent

ultrapure(GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 95.7%; Score 686.4; DB 6; Length 788;
Best Local Similarity 99.7%; Pred. No. 4.6e-160;
Matches 698; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GCTCACATACTAGTAGCCCTCCCTTTTCCTCGTGGCCACGATCTGCGAGGCA 60
Db |||||
Qy 90 GCTCACCATACTAGTAGCCCTCCCTTTTCCTCGTGGCCACGATCTGCGAGGCA 149
Db |||||

Qy 61 GCAGTGGGAACCTCAAGGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120
Db |||||

Qy 150 GCAGTGGGAACCTCAAGGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 209
Db |||||

Qy 121 GCCCTCGAGCAACATCTCATGCGAGAAGATCCAACTGACGAGATTCTATGAACGGGA 180
Db |||||

Qy 210 GCCCTCGAGCAACATCTCATGCGAGAAGATCCAACTGACGAGATTCTATGAACGGGA 269
Db |||||

Qy 181 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGG 240
Db |||||

Qy 270 CCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGG 329
Db |||||

Qy 241 ATCCTCTCAGCACCAAGAGAGGTGTCATAGCTGAACGATTGTGAGAACCAACCAAG 300
Db |||||

Qy 330 ATCCTCTCAGCACCAAGAGAGGTGTCATAGCTGAACGATTGTGAGAACCAACCAAG 389
Db |||||

Qy 301 GTGCATGTGCGAGGCATTGCAACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGGAG 360
Db |||||

Qy 390 GTGCATGTGCGAGGCATTGCAACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGGAG 449
Db |||||

Qy 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTTCAACAGTCGGCGCTTAG 420
Db |||||

Qy 450 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTTCAACAGTCGGCGCTTAG 509
Db |||||

Qy 421 GGCACACACGCTTGGACCTTGGACCTCGAAAGTGGCGGCGAGAGAGATCTAAACACC 480
Db |||||

Qy 510 GGCACACACGCTTGGACCTTGGACCTCGAAAGTGGCGGCGAGAGAGATCTAAACACC 569
Db |||||

Qy 481 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAGCTATTATCTATG 540
Db |||||

Qy 570 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAGCTATTATCTATG 629
Db |||||

Qy 541 GTTATGTTAGTTTGGTAAATAAAGATCATCATATATAGTGTGTTGATCGTGTTA 600
Db |||||

Qy 630 GTTATG-TTATGTTTGGTAAATAAAGATCATCATATATAGTGTGTTGATCGTGTTA 688
Db |||||

Qy 601 ACTAAGCGAAGCTTAGTTATATAGCACCTTTAGAGTGCCTTTATGCGGTTGTCTATCT 660
Db |||||

Qy 689 ACTAAGCGAAGCTTAGTTATATAGCACCTTTAGAGTGCCTTTATGCGGTTGTCTATCT 748
Db |||||

Qy 661 TTTGTTGCTGCGAGGTTGTAACCATCTTGAAATAATATAA 700
Db |||||

Qy 749 TTTGTTGCTGCGAAGTTGTAACCATCTTGAAATAATATAA 788
Db |||||

RESULT 2

CD038213
LOCUS
DEFINITION
UTPPI004 A09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI004_A09 5', mRNA sequence.
CD038213
ACCESSION
CD038213.1 GI:30420051

KEYWORDS

SOURCE

ORGANISM

EST.
Arachis hypogaea (peanut)

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE
1 (bases 1 to 689)

Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.

and Lynch, R. E. Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

Contact: Baozhu Guo

Molecular Genetics

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

FEATURES

source

Location/Qualifiers

1..689

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPP1004_A09"

/tissue_type="Immature pods"

/dev_stage="R6"

/lab_host="XLI-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library

(UTPP)"

/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;

cDNA library was constructed from peanut cultivar A13;

(NCV11XAR4). A13 has resistance to Aspergillus infection

and drought tolerance. The immature pods that developed to

R6 stage were collected from different plants, and placed

into liquid N2 immediately and stored in -80oC freezer.

Total RNA was isolated with TRIZOL-Reagent

ultrapure (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library

construction followed the protocol of by ZAP-cDNA Gigapack

III Gold cloning kit (Stratagene). The cDNA above 500bp

were collected after size-fraction. The inserts were

directionally cloned into Uni-ZAP XR vector using XhoI

EcoRI sites adapters. The lambda library was packed into

phages using Gigapack III Gold (Stratagene). The

un-amplified library was used to excise pBluescript

phagemids from the Uni-ZAP XR vector, and the phagemids

was used to transform the host bacteria SOLR. The library

was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 92.9%; Score 666; DB 6; Length 689;
Best Local Similarity 99.6%; Pred. No. 5.4e-155;
Matches 677; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 38 GCTGCCCGCATCTGCGAGCGAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGC 97
Db |||||

Qy 1 GCTGCCCGCATCTGCGAGCGAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGC 60
Db |||||

Qy 98 CAGCTCGAGAGGGCGAACTCTGAGGCCCTCGAGCAACATCTCATGAGAAGATCCAACT 157
Db |||||

Qy 61 CAGCTCGAGAGGGCGAACTCTGAGGCCCTCGAGCAACATCTCATGAGAAGATCCAACT 120
Db |||||

Qy 158 GACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGT 217
Db |||||

Qy 121 GACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGT 180
Db |||||

Qy 218 CCATATGATCGGAGAGCGCTGATCCTCTCAGCACCAAGAGAGGTGTTGCAATCAGCTG 277
Db |||||

Qy 181 CCATATGATCGGAGAGCGCTGATCCTCTCAGCACCAAGAGAGGTGTTGCAATCAGCTG 240
Db |||||

```
QY 278 AACGAGTTTGAGAACACCAAGGTGCATGTGCGAGGCGATTGCAACAGATCATGGAGAAC 337
Db 241 AACGAGTTTGAGAACACCAAGGTGCATGTGCGAGGCGATTGCAACAGATCATGGAGAAC 300
QY 338 CAGAGCGATAGTTGCGAGGAGGCAACAGGAGCAACAGATTCAAGAGGAGCTCAGGAAC 397
Db 301 CAGAGCGATAGTTGCGAGGAGGCAACAGGAGCAACAGATTCAAGAGGAGCTCAGGAAC 360
QY 398 TTGCCTCAACAGTGCAGGCTTAGGGACACACAGGTTGCGACTTGGACGTGCAAGTGGC 457
Db 361 TTGCCTCAACAGTGCAGGCTTAGGGACACACAGGTTGCGACTTGGACGTGCAAGTGGC 420
QY 458 GGCAGACAGACATTAACACCTTCTCAAAAAAGAAAAGAAAAGAAAAGAAAATAGC 517
Db 421 GGCAGACAGACATTAACACCTTCTCAAAAAAGAAAAGAAAAGAAAAGAAAATAGC 480
QY 518 TTATATATAGCTATTATCTATGTTATGTTTATGTTTGGTAAATAATAAAGATCATCACT 577
Db 481 TTATATATAGCTATTATCTATGTTATGTTTATGTTTGGTAAATAATAAAGATCATCACT 539
QY 578 ATATGAATGTTGATCGTGTAACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAG 637
Db 540 ATATGAATGTTGATCGTGTAACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAG 599
QY 638 TGCTTTTATGCGGTGTCATGTTTGTCTGTCGAGAGTTGTAAACCATCTTGAATAATA 697
Db 600 TGCTTTTATGCGGTGTCATGTTTGTCTGTCGAGAGTTGTAAACCATCTTGAATAATA 659
QY 698 TAAAAAGATCATGTTTGT 717
Db 660 TAAAAAGATCATGTTTGT 679
```

```
RESULT 3
LOCUS CD038384
DEFINITION UTPPI006_E02 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI006_E02 5', mRNA sequence.
ACCESSION CD038384
VERSION CD038384.1 GI:30420222
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 732)
Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
and Lynch, R. E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
FEATURES
source
1..732
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI006_E02"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/note="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
```

cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

```
Query Match 77.4%; Score 555.2; DB 6; Length 732;
Best Local Similarity 93.1%; Pred. No. 2.2e-127;
Matches 617; Conservative 0; Mismatches 9; Indels 37; Gaps 2;

QY 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCCGCCATCTCGCGAGCA 60
Db 70 GCTCACCATACTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCCGCCATCTCGCGAGCA 129
QY 61 GCAGTGGGAATCCAGAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 120
Db 130 GCAGTGGGAATCCAGAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 189
QY 121 GCGCTGCGAGCAACATCTCATGCGAAGATCCAAAGTGAACGAGGATTCATATGAACGGGA 180
Db 190 GCGCTGCGAGCAACATCTCATGCGAAGATCCAAAGTGAACGAGGATTCATATGAACGGGA 249
QY 181 CCGGTACAGCCCTAGTCAG-----CATCC 204
Db 250 CCGGTACAGCCCTAGTCAG-----CATCC 309
QY 205 GTACAGCCCTAGTCATATGATCGGAGAGCGCTGGATCTCTCAGCACCACAGAGAGGTG 264
Db 310 GTACAGCCCTAGTCATATGATCGGAGAGCGCTGGATCTCTCAGCACCACAGAGAGGTG 369
QY 265 TTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGCATGTGCGAGGCAATTCGAACA 324
Db 370 TTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGCATGTGCGAGGCAATTCGAACA 429
QY 325 GATCATGGAGAACCGAGCGATAGTTGCGAGGAGGCAACAGGAGCAACAGTTCAAGAG 384
Db 430 GATAATGGAGAACCGAGCGATAGTTGCGAGGAGGCAACAGGAGCAACAGTTCAAGAG 489
QY 385 GGAGCTCAGGAACTTGCCTCAACAGTGCAGGCGCTTAGGGCACCACAGCGTTGCGACTTGA 444
Db 490 GGAGCTCAGGAACTTGCCTCAACAGTGCAGGCGCTTAGGGCACCACAGCGTTGCGACTTGA 549
QY 445 COTCGAAAGTGGCGCAGACAGATATCTAAACACTATCTCAAAAAAGAAAAGAAAAG 504
Db 550 AGTCGAAAGTGGCGCAGACAGATATCTAAACACTATCTCAAAAAAGAAAAGAAAAG 609
QY 505 AAAAGAAATAGCTTATATATAGCTTATCTATGTTTATGTTTATGTTTATGTTTATGTT 564
Db 610 AAAAGAAATAGCTTATATATAGCTTATCTATGTTTATGTTTATGTTTATGTTTATGTT 668
QY 565 AAAGATCATCACTATATGAAATGTTGATCGGTGTTAACTAAGCGAAGCTTAGGTTATATG 624
Db 669 AAAGATCATCACTATATGAAATGTTGATCGGTGTTAACTAAGCGAAGCTTAGGTTATATG 728
QY 625 AGC 627
Db 729 AGC 731
```



```

QY 197 CAG-----GATCGTACAGCCCTAGTCCA 220
Db 61 CAGGATCCGTACAGCCCTAGTACAGACCCGACAGAGCGTATCGTACAGCCCTAGTCCA 120
QY 221 TATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGGTGTTCATAGAGCTGAAC 280
Db 121 TATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGGTGTTCATAGAGCTGAAC 180
QY 281 GAGTTTGAGAACCAACAAAGTGCATGTGCGAGGCATTCACACAGATCATGGAGAACCCAG 340
Db 181 GAGTTTGAGAACCAACAAAGTGCATGTGCGAGGCATTCACACAGATCATGGAGAACCCAG 240
QY 341 AGCGATAGTTGCGAGGGGCGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTG 400
Db 241 AGCGATAGTTGCGAGGGGCGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTG 300
QY 401 CCTCAACAGTGGCGGCTTAGGGCACACAGCGTTGCGACTTGCAGCTGCGAAAGTGGCGGC 460
Db 301 CCTCAACAGTGGCGGCTTAGGGCACACAGCGTTGCGACTTGCAGCTGCGAAAGTGGCGGC 360
QY 461 AGAGACAGATCTAAACACCTATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGTACTTA 520
Db 361 AGAGACAGATCTAAACACCTATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGTACTTA 420
QY 521 TATATAAGCTTATATCTATGTTATGTTTATGTTTGTGTAATAAAGATCATCATATA 580
Db 421 TATATAAGCTTATATCTATGTTATGTTTATGTTTGTGTAATAAAGATCATCATATA 479
QY 581 TGAATGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATAGGACCTTTAGAGTGC 640
Db 480 TGAATGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATAGGACCTTTAGAGTGC 539
QY 641 TTTT--ATGCGGTGTCTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAATAT 698
Db 540 TTTTATGCGGTGTCTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAATAT 599
QY 699 AAAAAATCATGTTTGT 717
Db 600 NGAAGATCANGGTTTGT 618

RESULT 6
CD038306
LOCUS
DEFINITION
  UTPI005_D03 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
  Arachis hypogaea cDNA clone UTPI005_D03 5', mRNA sequence.
ACCESSION
  CD038306
VERSION
  CD038306.1 GI:30420144
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
  1 (bases 1 to 405)
  Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
  and Lynch,R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3.
  Location/Qualifiers
    1..405
      /organism="Arachis hypogaea"
      /mol_type="mRNA"

FEATURES
  source

```

```

/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPI005_D03"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). cDNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match 43.9%; Score 315; DB 6; Length 405;
Best Local Similarity 99.7%; Pred. No. 1.4e-67;
Matches 326; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 391 CAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCACACAGCGTTGCGACTTGGACGTCGA 450
Db 77 CAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCACACAGCGTTGCGACTTGGACGTCGA 136
QY 451 AAGTGGCGGCGAGACAGATACATAACACCTATCTCAAAAAAGAAAGAAAGAAAGA 510
Db 137 AAGTGGCGGCGAGACAGATACATAACACCTATCTCAAAAAAGAAAGAAAGAAAGA 196
QY 511 AATAGCTTATATATAAGCTATTATCTATGTTTATGTTTGTGTAATAATAAGAT 570
Db 197 AATAGCTTATATATAAGCTATTATCTATGTTTATGTTTGTGTAATAATAAGAT 255
QY 571 CATCATATATCAATGCTGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATGACACC 630
Db 256 CATCATATATCAATGCTGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATGACACC 315
QY 631 TTTAGAGTCTTTTATGCGGTTGCTATGTTTGTGTCGAGAGTTGTAACCATCTTGA 690
Db 316 TTTAGAGTCTTTTATGCGGTTGCTATGTTTGTGTCGAGAGTTGTAACCATCTTGA 375
QY 691 AATAATATAAAGATCATGTTTGT 717
Db 376 AATAATATAAAGATCATGTTTGT 402

RESULT 7
CD038459
LOCUS
DEFINITION
  UTPI007_F09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
  Arachis hypogaea cDNA clone UTPI007_F09 5', mRNA sequence.
ACCESSION
  CD038459
VERSION
  CD038459.1 GI:30420297
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
  1 (bases 1 to 308)

```

```

AUTHORS      Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
              and Lynch,R.E.
TITLE        Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
FEATURES     Location/Qualifiers
source      1..308
            /organism="Arachis hypogaea"
            /mol_type="mRNA"
            /cultivar="Al3"
            /db_xref="taxon:3818"
            /clone="UTPPI007_F09"
            /tissue_type="Immature pods"
            /dev_stage="R6"
            /lab_host="Xl1-blue"
            /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
            (UTPP)"
            /note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
            cDNA library was constructed from peanut cultivar Al3
            (NCV11XR4). Al3 has resistance to Aspergillus infection
            and drought tolerance. The immature pods that developed to
            R6 stage were collected from different plants, and placed
            into liquid N2 immediately and stored in -80oC freezer.
            Total RNA was isolated with TRIzol-Reagent
            ultrapure (GIBCOBRL). mRNA was extracted and purified from
            total RNA (Promega). cDNA synthesis and library
            construction followed the protocol of by ZAP-cDNA Gigapack
            III Gold cloning kit (Stratagene). The cDNA above 500bp
            were collected after size-fraction. The inserts were
            directionally cloned into Uni-ZAP XR vector using XhoI
            EcoRI sites adapters. The lambda library was packed into
            phages using Gigapack III Gold (Stratagene). The
            un-amplified library was used to excise pbluescript
            phagemids from the Uni-ZAP XR vector, and the phagemids
            was used to transform the host bacteria SOLR. The library
            was constructed by Dr. Meng Luo and Dr. Phat Dang."
ORIGIN
Query Match      40.3%; Score 289; DB 6; Length 308;
Best Local Similarity 99.7%; Pred. No. 4.1e-61;
Matches 300; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 417 TTAGGCGCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGATACTAA 476
DB 1 TTAGGCGCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGATACTAA 60
QY 477 CACCTATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATC 536
DB 61 CACCTATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGTATATATAAGCTATTATC 120
QY 537 TATGGTTATGTTTGTAGTTTGGTAAATAATAAGATCATCACTATGTAATGTGTGATCGT 596
DB 121 TATGGTTATG- TTAGTTTGGTAAATAATAAGATCATCACTATGTAATGTGTGATCGT 179
QY 597 GTTAACCTAAGGCAAGCTTAGGTTATATAGCACCTTTTAGAGTGCTTTTATGGCGTTGTCT 656
DB 180 GTTAACCTAAGGCAAGCTTAGGTTATATAGCACCTTTTAGAGTGCTTTTATGGCGTTGTCT 239
QY 657 ATGTTTGTGCTGCGAGAGTTGTAACCATCTTGAATAATATAATAAAGATCATGTTTGT 716
DB 240 ATGTTTGTGCTGCGAGAGTTGTAACCATCTTGAATAATATAATAAAGATCATGTTTGT 299
QY 717 T 717
DB 300 T 300

```

```

RESULT 8
CO897504
LOCUS       CO897504
DEFINITION EST00009 Peanut Lambda Express library Arachis hypogaea cDNA 5',
            mRNA sequence.
ACCESSION  CO897504
VERSION    CO897504.1 GI:51237294
KEYWORDS   EST.
SOURCE     Arachis hypogaea (peanut)
ORGANISM   Arachis hypogaea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
            Aeschynomeneae; Arachis.
REFERENCE  1 (bases 1 to 375)
            Yan,Y.S., Wang,L. and Huang,S.Z.
            Expressed sequence tags of mid-matured cotyledons of peanut
            Unpublished (2004)
            Contact: ShangZhi Huang
            Plant Development and Molecular Biology
            Sun Yat-sen University, Department of Biology and Biotechnology
            Guangzhou, 510275, China
            Tel: (860)02084036592
            Email: yys9803@yahoo.com.cn
            Seq primer: pTriplex2 Forward.
FEATURES   Location/Qualifiers
source     1..375
            /organism="Arachis hypogaea"
            /mol_type="mRNA"
            /cultivar="Yueyou 523"
            /db_xref="taxon:3818"
            /tissue_type="Cotyledons"
            /dev_stage="Mid-matured stage"
            /lab_host="E.coli BM25.8"
            /clone_lib="Peanut Lambda Express library"
            /note="Organ: Seed; Vector: lambdaTriplex2"
ORIGIN
Query Match      39.5%; Score 283.4; DB 7; Length 375;
Best Local Similarity 88.8%; Pred. No. 1e-59;
Matches 333; Conservative 0; Mismatches 6; Indels 36; Gaps 1;
QY 103 CGAGAGGGCGAACCCTGAGGCCCTGCGAGCAACATCTCATGCAGAAAGATCCAAACGTGACGA 162
DB 1 CGAGAGGGCGAACCCTTAGGCCCTGCGAGCAACATCTCATGCAGAAATCCAAACGTGACGA 60
QY 163 GGATTATATGAACCGGACCCGCTACAGCCCTAGTCAG----- 199
DB 61 GGATTATATGACGGGACCCGCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCAGGA 120
QY 200 -----GATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGAATCCTC 246
DB 121 CCCGACACAGAGTGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGAATCCTC 180
QY 247 TCAGCACCACAGAGAGGTGTTGCAATGAGCTGAACAGAGTTTGAGAAACAACCAAGGTGCAT 306
DB 181 TCAGCACCAGGAGAGGTGTTGCAATGAGCTGAACAGAGTTTGAGAAACAACCAAGGTGCAT 240
QY 307 GTGCGAGGCAATTGCAACAGATCATGGAGAACACAGAGCGATAGTTGCGAGGGGAGGCAACA 366
DB 241 GTGCGAGGCAATTGCAACAGATAATGGAGAACACAGAGCGATAGTTGCGAGGGGAGGCAACA 300
QY 367 GGAGCAACAGATTCAAGAGGGAGCTCAGGAACCTTGCTCTCAACAGTCGCGCTTAGGGCACC 426
DB 301 GGAGCAACAGATTCAAGAGGGAGCTCAGGAACCTTGCTCTCAACAGTCGCGCTTAGGGCACC 360
QY 427 ACAGCGTTGCGACTT 441
DB 361 ACAGCGTTGCGACTT 375

```

```

RESULT 9

```

```

CD038415      696 bp      mRNA      linear      EST 07-MAY-2003
LOCUS      UTPPI006_H11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION      Arachis hypogaea cDNA clone UTPPI006_H11 5', mRNA sequence.
ACCESSION      CD038415
VERSION      CD038415.1 GI:30420253
KEYWORDS      EST.
SOURCE      Arachis hypogaea (peanut)
ORGANISM      Arachis hypogaea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE      1 (bases 1 to 696)
AUTHORS      Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
              and Lynch,R.E.
TITLE      Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
              Location/Qualifiers
FEATURES             source
                     1..696
                     /organism="Arachis hypogaea"
                     /mol_type="mRNA"
                     /cultivar="A13"
                     /db_xref="taxon:3818"
                     /tissue_type="Immature pods"
                     /dev_stage="R6"
                     /lab_host="XL1-blue"
                     /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
                     (UTPP)"
                     /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
                     cDNA library was constructed from peanut cultivar A13
                     (NCV11XAR4). A13 has resistance to Aspergillus infection
                     and drought tolerance. The immature pods that developed to
                     R6 stage were collected from different plants, and placed
                     into liquid N2 immediately and stored in -80oC freezer.
                     Total RNA was isolated with TRIZOL-Reagent
                     ultrapure (GIBCOBRL). mRNA was extracted and purified from
                     total RNA (Promega). cDNA synthesis and library
                     construction followed the protocol of by ZAP-cDNA Gigapack
                     III Gold cloning kit (Stratagene). The cDNA above 500bp
                     were collected after size-fraction. The inserts were
                     directionally cloned into Uni-ZAP XR vector using XhoI
                     EcoRI sites adapters. The lambda library was packed into
                     phages using Gigapack III Gold (Stratagene). The
                     un-amplified library was used to excise pBluescript
                     phagemids from the Uni-ZAP XR vector, and the phagemids
                     was used to transform the host bacteria SOLR. The library
                     was constructed by Dr. Meng Luo and Dr. Phat Dang."
ORIGIN
Query Match      31.2%; Score 223.8; DB 6; Length 696;
Best Local Similarity 71.9%; Pred. No. 8.1e-45;
Matches 346; Conservative 0; Mismatches 107; Indels 28; Gaps 3;

QY      219      CATATGATCGGAGGCGGTGGATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGA 278
Db      234      CCTACGATATTAGGAGTACTCGATCTCCGACCAGCACAGAGGTCTGGATGACTGA 293
QY      279      ACAGTTTGAACAACCAAGGTGATGTGCGAGGCAATTGCAACAGATCATGGAAC 338
Db      294      ACCAGATGAGAAACACACAGAGATGATGTGCGAGGCAATTGCGAGCAGATAATGGAAC 353
QY      339      AGAGCCATAGTTGACGGGAGGCAACAGGACCAAGTTCAGAGGGAGCTCAGAACT 398

```

```

Db      354      AGTGCATAGGTTGCAGGACAGGCAAAATGGTCAGCAGTTCAAGAGAGCTCATGAAC 413
QY      399      TGCCTCAACAGTGGCGGCTTAGGGACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCG 458
Db      414      TGCCCCAACAGTGTAACTTTTAGGGCACCACAGCGTTGCGATTGGACGT-----GA 464
QY      459      GCAGAGACAGACTAAACACACCTATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGT 518
Db      465      GTGGCGGCGAGATGCTAGACTCAAAAATAATAATCTGTGCCAAAACAAACTAGTAGGAAGT 524
QY      519      TATATATAGCTATTATCTATGTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 578
Db      525      AGCTTATGAGCTATTATGTTATGTT-----GTTTCGTTTATATAAACATCATCACTG 577
QY      579      TATGAATGTTGATGCTGTTAACTAAAGCAAGCTTAGGTTATATATAGCACCTTTAGAGT 638
Db      578      TATGAATGTTGTCAT-----AGCTAGGTAAGGTTATATGAGCCCTTCGGTGT 625
QY      639      GCTTTATGCGCTGCTCTATGTTTGTCTCCAGAGTTGTAACCATCTTGAATAATAT 698
Db      626      GCTCTTATGGCTTTACCTGTGTTTGTCTACTGCAAGTTTAAACCAACCATGAATAAAG 685
QY      699      A 699
Db      686      A 686

RESULT 10
CD038394
LOCUS      UTPPI006_F02 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION      Arachis hypogaea cDNA clone UTPPI006_F02 5', mRNA sequence.
ACCESSION      CD038394
VERSION      CD038394.1 GI:30420232
KEYWORDS      EST.
SOURCE      Arachis hypogaea (peanut)
ORGANISM      Arachis hypogaea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE      1 (bases 1 to 732)
AUTHORS      Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
              and Lynch,R.E.
TITLE      Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
              Location/Qualifiers
FEATURES             source
                     1..732
                     /organism="Arachis hypogaea"
                     /mol_type="mRNA"
                     /cultivar="A13"
                     /db_xref="taxon:3818"
                     /clone="UTPP1006_F02"
                     /tissue_type="Immature pods"
                     /dev_stage="R6"
                     /lab_host="XL1-blue"
                     /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
                     (UTPP)"
                     /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
                     cDNA library was constructed from peanut cultivar A13
                     (NCV11XAR4). A13 has resistance to Aspergillus infection
                     and drought tolerance. The immature pods that developed to
                     R6 stage were collected from different plants, and placed
                     into liquid N2 immediately and stored in -80oC freezer.
                     Total RNA was isolated with TRIZOL-Reagent

```

ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 30.0%; Score 215.2; DB 6; Length 732;
Best Local Similarity 72.0%; Pred. No. 1.1e-42;
Matches 347; Conservative 0; Mismatches 108; Indels 27; Gaps 4;
QY 219 CATATGATCGGAGGCGGTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGA 278
DB 262 CCTACGATATTAGGAGTACTCGATCCTCCGACCAGCAAGAGGTGTCGATGAGCTGA 321
QY 279 ACGAGTTTCAGAACCAACCAAGGTGATGTCGAGGCATTCGACAGATCATGGAGAACC 338
DB 322 ACGAGATGAGAACACACAGAGATGCATGTGCGAGGCATTCGACGAGATAATGGAGAACC 381
QY 339 AGAGCGATAGGTTGACGGGAGGCAACAGGCAACAGTTCAGAGGGAGCTCAGGAACCT 398
DB 382 AGTGCATAGGTTGAGGACAGCAAAATGTCGACGATTCAGAGAGAGCTCATGAACCT 441
QY 399 TGCCTCAACAGTGCAGGCTTAGGGCACCAAGCGTTGCGACTTTGACGTCGAAAGTGGCG 458
DB 442 TGCCCCAACAGGTAACTTTAGGGCACCAAGCGTTGCGATTTGGACGT---GAGTGGCG 498
QY 459 GCAGACAGATCACTAAACACCTATCTCAAAAAGAAAGAAAGAAAGAAATAGCT 518
DB 499 GCAGATGCTAGACTCAAAAATAATATCTGTGCCAAAACAAACTTAGTAGGAAGTAGCTT 558
QY 519 TATATATAGCTATTATCTATGTTATGTTTATGTTTGTGTAATAATAAGATCATCACTA 578
DB 559 ATATTATAGCTATTATGATGCTT-----GTTTCGTTAATAATAATATCATCACTG 611
QY 579 TATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATGAGCACCTTTTAGAGT 638
DB 612 TATGAATGTTGATGAT-----AGGTAAGGTTATATGAGCACCTTCGGTGT 655
QY 639 GCTTTTATGCG-GTTGTCATGTTTGTGTCGAGAGTTGTAACCATCTTGAATAATA 697
DB 656 GCTCTTATGGCTTTTACCTATGTTTGTCTACTGCAAAAGTTTTAACCATGAATAAAA 715
QY 698 TA 699
DB 716 GA 717

RESULT 11
CD038326
LOCUS CD038326.1 GI:30420164
DEFINITION UTPI005 F10 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPI005_F10 5', mRNA sequence.
ACCESSION CD038326
VERSION CD038326.1
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 661)
REFERENCE Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea

JOURNAL COMMENT

Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: n3

FEATURES

source location/Qualifiers
1..661
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTPI005_F10"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar Al3 (NCV11XAR4). Al3 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 29.1%; Score 208.8; DB 6; Length 661;
Best Local Similarity 65.1%; Pred. No. 4.4e-41;
Matches 457; Conservative 0; Mismatches 164; Indels 81; Gaps 7;
QY 4 CACCACTACTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACCGCATCTGC---GAGGCA 60
DB 25 CACCATCTGTGTAGCCCTCTTGGCCCTCGTCTGTGGCACACGCTCCGCGATGAGCGG 84
QY 61 GCAGTGGGAACTCCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACT 117
DB 85 CGAGAGGGGGCGACNAGGGGACTCATCAAGCTGCGAGAGGGCAGGTAGACAGGTTGAACCT 144
QY 118 GAGGCCCTCGGAGCAACATCTCATGCAAGAATCCACGTGACGAGGATTCATATGAACG 177
DB 145 CAAGCCCTCGGAGCAGCACATAATGCAGAGGATAATGGCGAGCAAGACGAGTACGACT- 203
QY 178 GGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGCGC 237
DB 204 -----CCTACGATATTAGGAGTAC 222
QY 238 TGGATCTCTTCAGCACCAAGAGAGGTTGCAATGAGCTGAACAGTTTGAACAACA 297
DB 223 TCGATCTCTCCAGCCAGCAACAGAGGTGCTGCGATGAGCTGGACCAAGATGGAGNACACAGA 282
QY 298 AAGGTGTCATGTCCGAGGCATTTGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCAGGG 357
DB 283 GAGATGTCATGTGCGAGGCATTTGCAGCAGATAATGAGAAACCAAGTCCGATAGGTTGCAGGA 342
QY 358 GAGGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCT 417
DB 343 CAGGCAAAATGGTGCAGCAGTTTCAAGAGAGAGCTCATGAACCTTGCCCTCAACAGTGAACCT 402


```

QY 418 TAGGGACACAGCGTTGGCACTTGGAGCTGGAAGTGGCGGAGAGACAGATACTAAAC 477
Db 403 CAGGGACACAGCGTTGGCACTTGGAGCTGGAAGTGGCGGAGAGATGCTAGAC 453
QY 478 ACCTATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGTATATATATAAGCTTATTATCT 537
Db 454 TCAAAATATATATCTGTGCCAAAGAAAGAAAGTATAGGAAGTCTATAGAGCTTATTATGT 513
QY 538 ATGGTTAAGTTTATGTTTGGTAATAATAAAGATCATCATATATATGAATGTGTTGATCGTG 597
Db 514 ATGCTT-----GTTTCGTTAATAATAATAATATCATCATGTATGAATGTGGTGAT---- 562
QY 598 TTAACCTAAGCAGCTTAGCTTATATAGCAGCACCTTTAGAGTGCTTTTATGGCGCTTGCTTA 657
Db 563 -----AGGTAAGGTTTATAGGACCACTTGGTGTCTTATGGCTTTACCTA 609
QY 658 TGTTCCTGCTGAGAGTTGTAAACCATCTTGAATAATAATA 699
Db 610 TGTTTGTCTACTGCAAGTTTANCCACACCTGAAATAAAGA 651

RESULT 12
CO897508 211 bp mRNA linear EST 13-AUG-2004
LOCUS EST000013 Peanut Lambda Express library Arachis hypogaea cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CO897508
VERSION CO897508.1 GI:51237298
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 211)
AUTHORS Yan, Y.S., Wang, L. and Huang, S.Z.
TITLE Expressed sequence tags of mid-matured cotyledons of peanut
JOURNAL Unpublished (2004)
COMMENT Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (860)02084036592
Email: yys9803@yahoo.com.cn
Seq primer: pTriplex2 Forward.
Location/Qualifiers
FEATURES
source
1..211
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/tissue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/notes="Organ: Seed; Vector: lambdaTriplex2"

ORIGIN
Query Match 20.0%; Score 143.4; DB 7; Length 211;
Best Local Similarity 83.94; Pred. No. 7.3e-25;
Matches 162; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 267 GCAATGAGCTGAACGAGTTTGAACAACCAAGGTCATGTCGAGGCAATGCAACAGA 326
Db 11 GCATGAGCTGAACGAGATGAGAACACACAGAGATGATGTCGAGGCAATGTCAGCAGA 70
QY 327 TCATGGAGAACCGAGCGATAGGTTGCAGGGGAGGCAACAGGACCAAGTTCAAGAGGG 386
Db 71 TAATGAGAACCAAGTCGATAGGTTGCAGGACGCAAGCAATGGTGCAGCAGTTCAAGAGAG 130
QY 387 AGCTCAGGAACCTTGCCTCAACAGTCGGCGCTTAGGACCACACAGCGTTGCGACTTGGACG 446

```

```

Db 131 AGCTCATGAACCTGCCCCAACAGTGTAACCTTTAGGACACACAGCGTTGCGATTGGACG 190
QY 447 TCGAAAGTGGCGG 459
Db 191 TGAGTGGCGGCGAG 203

RESULT 13
CO897506 286 bp mRNA linear EST 13-AUG-2004
LOCUS EST000011 Peanut Lambda Express library Arachis hypogaea cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CO897506
VERSION CO897506.1 GI:51237296
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 286)
AUTHORS Yan, Y.S., Wang, L. and Huang, S.Z.
TITLE Expressed sequence tags of mid-matured cotyledons of peanut
JOURNAL Unpublished (2004)
COMMENT Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (860)02084036592
Email: yys9803@yahoo.com.cn
Seq primer: pTriplex2 Forward.
Location/Qualifiers
FEATURES
source
1..286
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/tissue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/notes="Organ: Seed; Vector: lambdaTriplex2"

ORIGIN
Query Match 19.2%; Score 137.6; DB 7; Length 286;
Best Local Similarity 80.5%; Pred. No. 2.1e-23;
Matches 161; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 219 CATATGATCGAGAGCGCTGGATCCTCTCAGACCAAGAGAGGTTTGAATGAGCTGA 278
Db 87 CCTACGATATTAGGAGTACTCGATCCTCGGACCAAGAGAGGTTCTGCGATGAGCTGA 146
QY 279 ACAGTTTTCAGAACCAACCAAGGTCGATGTCGAGGCAATGCAACAGATCATGGAGAACC 338
Db 147 ACAGATGAGAAACACACAGAGATGATGTCGAGGCAATGTCGAGAGATAATGGAGAACC 206
QY 339 AGAGCGATAGGTTGCGAGGGGAGGCAACAGGACCAAGTTCGAAGGGAGCTCAGGAAC 398
Db 207 ATGCGATAGGTTGCGAGGACAGGCAAAATGTCGAGGAGTTCAAGAGAGAGCTCATGACT 266
QY 399 TGCCTCAACAGTGGCGCTTT 418
Db 267 TGCCCCAACAGTGTAACTTT 286

RESULT 14
CD038092 680 bp mRNA linear EST 07-MAY-2003
LOCUS UTPI002 D04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION Arachis hypogaea cDNA clone UTPI002_D04 5', mRNA sequence.
ACCESSION CD038092

```

```

VERSION      CD038092.1  GI:30419930
KEYWORDS     EST.
SOURCE       Arachis hypogaea (peanut)
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE    1 (bases 1 to 680)
AUTHORS      Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
              and Lynch, R.E.
TITLE        Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
FEATURES     Location/Qualifiers
             1..680
             /organism="Arachis hypogaea"
             /mol_type="mRNA"
             /cultivar="A13"
             /db_xref="taxon:3818"
             /clone="UTPPI002_D04"
             /tissue_type="Immature pods"
             /dev_stage="R6"
             /lab_host="XL1-blue"
             /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
             (UTPP)"
             /note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
             cDNA library was constructed from peanut cultivar A13;
             (NCV1XA4). A13 has resistance to Aspergillus infection
             and drought tolerance. The immature pods that developed to
             R6 stage were collected from different plants, and placed
             into liquid N2 immediately and stored in -80oC freezer.
             Total RNA was isolated with TRIzol-Reagent
             ultrapur (GIBCOBRL). mRNA was extracted and purified from
             total RNA (Promega). cDNA synthesis and library
             construction followed the protocol of by ZAP-cDNA Gigapack
             III Gold cloning kit (Stratagene). The cDNA above 500bp
             were collected after size-fraction. The inserts were
             directionally cloned into Uni-ZAP XR vector using XhoI
             EcoRI sites adapters. The lambda library was packed into
             phages using Gigapack III Gold (Stratagene). The
             un-amplified library was used to excise phagescript
             phagemids from the Uni-ZAP XR vector, and the phagemids
             was used to transform the host bacteria SOLR. The library
             was constructed by Dr. Meng Luo and Dr. Phat Dang."
             source
             1..351
             /organism="Arachis hypogaea"
             /mol_type="mRNA"
             /cultivar="Yueyou 523"
             /db_xref="taxon:3818"
             /tissue_type="Cotyledons"
             /dev_stage="Mid-matured stage"
             /lab_host="E.coli BM25.8"
             /clone_lib="Peanut Lambda Express library"
             /note="Organ: Seed; Vector: lambdaTriplex2"
ORIGIN
Query Match      18.6%; Score 133.2; DB 6; Length 680;
Best Local Similarity 58.3%; Pred. No. 3e-22;
Matches 344; Conservative 0; Mismatches 213; Indels 33; Gaps 5;

QY 76 AGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGGAACCTGAGGCCCTCGAGCAACA 135
Db 107 AGGGGATACCACTGTCAGAGGCAGTTGCGAGGGGCAACCTGAGGCCCTGTGAGGAACA 166
QY 136 TCTCATGCAAGAGATCCAAAGCTGACGAGGATTTCATATGAACGGACCCCGTACAGCCCTAG 195
Db 167 CATAGGCAAGGGTGGAGCAAGAGCAAGAGCAAGAGCAGATGCCCTACAGCCA 226
QY 196 TCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCAACA 255
Db 227 ACGGGGATCCAGGAGCAGCAGCACC-----GGCGAATCTGACGAAGATCAAGA 274
QY 256 AGAGAGGTTTGCATGAGCTGAACGAGTTTGAGACACACCAAGGTGATGTCGAGGC 315
Db 275 GCAAGGGTCTGCAACGAGCTCAACCGGTTTCAGAAATACCAAGGTGCATGTGCCAGGC 334

```

```

QY 316 ATTGCAACAGATCATGGAGAACCCAGAGCGATAGGTT---GCAGGGGAGGCAACAGGAGCA 372
Db 335 ACTTCAAAGATCTCCAGAACCCAGAGCTTTAGGTTCCAGCAGGACAGGACCGAGTTCGA 394
QY 373 ACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGATGCGGCGCTTAGGGCAACCAAGCG 432
Db 395 TCAGATGGAGGGAGCTCAGGAACCTGCCCCATAAATCGCGGTTTCAGGTCAACCAAGCG 454
QY 433 TTGCGACATT-----GGACGTGMAAGTGGCGGCAGAGACAGATACATAACACC 480
Db 455 TTGCGACCTTAGTAGCCGCGACGCCCTACTAAACAGACATAGCACCTTTGGCTTTAATTGCG 514
QY 481 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGCTATTATCTATG 540
Db 515 TTACCCACACAGAGAAATAACATGATATAATAATAATAGCTTTTTTACAGCTATTCTA 574
QY 541 GTTATGTTTGGTTTGGTAAATAAAGATCATCATCTATAGA--ATGTTTGGATCGGTG 598
Db 575 TGTCTATGTTGTTTGGTAGCAATAAAGATCATCACCATTTTATGTAGTGGTATCGTAT 634
QY 599 TAACTAAGCAAGCTTAGGTTATATAGACACCTTTAGAGTGCCTTTATGG 648
Db 635 TGTCTGTGCGGAAGTTATAT-----GGGACACTTTAAATGTGCTTTTATGG 680

RESULT 15
CO897507
LOCUS      EST000012 Peanut Lambda Express library Arachis hypogaea cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CO897507
VERSION    CO897507.1  GI:51237297
KEYWORDS   EST.
SOURCE     Arachis hypogaea (peanut)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE  1 (bases 1 to 351)
AUTHORS    Yan, Y.S., Wang, L. and Huang, S.Z.
TITLE      Expressed sequence tags of mid-matured cotyledons of peanut
JOURNAL    Unpublished (2004)
COMMENT     Contact: Shangzhi Huang
              Plant Development and Molecular Biology
              Sun Yat-sen University, Department of Biology and Biotechnology
              Guangzhou, 510275, China
              Tel: (860)02084036592
              Email: YYS9803@yahoo.com.cn
              Seq primer: pTriplex2 Forward.
              Location/Qualifiers
              1..351
              /organism="Arachis hypogaea"
              /mol_type="mRNA"
              /cultivar="Yueyou 523"
              /db_xref="taxon:3818"
              /tissue_type="Cotyledons"
              /dev_stage="Mid-matured stage"
              /lab_host="E.coli BM25.8"
              /clone_lib="Peanut Lambda Express library"
              /note="Organ: Seed; Vector: lambdaTriplex2"
ORIGIN
Query Match      18.4%; Score 131.8; DB 7; Length 351;
Best Local Similarity 80.6%; Pred. No. 6.1e-22;
Matches 154; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 219 CATATGATCGGAGAGCGCTCGATCCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGA 278
Db 161 CCTAGATATTAGGAGTACTCGATCCTCCGACCACACAGAGGTGCTCGATGAGCTGG 220
QY 279 ACGAGTTTGAGAACCAACCAAGGTGTCATGTGCGAGGTCATTGCAACAGATCATGGAGAAC 338

```

Db	231	ACCAGATGGAGAACACAGAGATGCAATGTGGAGGCATTTGCAGCAGATAATGSGAACC	280
Qy	339	AGAGCGATAGTGTGCAGGGGAGGCCAACAGGAGCAACAGTTTCAAGAGGGAGCTTCAGGAAC	398
Db	281	AGTGCATAGTGTGCAGACAGGCCAAATGGTGCACAGTTTCAAGAGAGNGCTCATGAACT	340
Qy	399	TGCCTCAACAG	409
Db	341	TGCCTCAACAG	351

Search completed: August 24, 2005, 03:44:12
Job time : 1685.67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 117.962 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcaccatactagtagcc.....taaaagatcatgtttgttt 717

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	100.0	717	4	US-09-106-872A-20
2	717	100.0	717	4	US-09-191-593-6
3	717	100.0	717	4	US-09-191-593-9
4	713.8	99.6	743	4	US-09-106-872A-1
5	65.6	9.1	777	2	US-08-618-911-3
6	64	8.9	777	2	US-08-618-911-5
7	63.4	8.8	770	3	US-08-938-675A-1
8	63.4	8.8	770	4	US-09-531-727-1
9	58.8	8.2	7218	1	US-08-232-463-14
10	54.4	7.6	723	2	US-08-618-911-1
11	44.4	6.2	168104	4	US-09-949-016-12026
12	44.4	6.2	168105	4	US-09-949-016-16554
13	44.4	6.2	1664976	4	US-08-916-421B-1
14	44.4	6.2	1664976	4	US-09-692-570-1
15	44.2	6.2	56685	4	US-09-949-016-14026
16	44	6.1	31739	4	US-09-949-016-16226
17	43.6	6.1	601	4	US-09-949-016-44480
18	43.6	6.1	34855	4	US-09-949-016-13004
19	43.6	6.1	56302	4	US-09-949-016-11892
20	43.2	6.0	251672	4	US-09-949-016-17296
21	43.2	6.0	251682	4	US-09-949-016-11973
22	42.6	5.9	133613	4	US-09-949-016-15824
23	42.4	5.9	289	3	US-09-007-005-17
24	42.4	5.9	289	3	US-09-244-796-17
25	42.2	5.9	95890	4	US-09-949-016-16412
26	42	5.9	5630	2	US-08-937-931-1
27	42	5.9	5630	3	US-09-285-502-1

ALIGNMENTS

RESULT 1

US-09-106-872A-20
; Sequence 20, Application US/09106872A

; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannion, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Atachis hypogaea
US-09-106-872A-20

Query Match 100.0%; Score 717; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.9e-189;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCAGCATCTCGGAGGCA	60
Db	1	GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCAGCATCTCGGAGGCA	60
QY	61	GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCACTGAG	120
Db	61	GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCACTGAG	120
QY	121	GCCTCGCAGGCAACATCTCATGAGAGATCCAACTGACGAGGATTCATATGAAACGGGA	180
Db	121	GCCTCGCAGGCAACATCTCATGAGAGATCCAACTGACGAGGATTCATATGAAACGGGA	180
QY	181	CCGCTACAGCCCTAGTCAGGATCCGCTACAGCCCTAGTCATATGATCGGAGGCGCTGG	240
Db	181	CCGCTACAGCCCTAGTCAGGATCCGCTACAGCCCTAGTCATATGATCGGAGGCGCTGG	240
QY	241	ATCCTCTCAGCACCAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG	300
Db	241	ATCCTCTCAGCACCAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG	300

301	Qy	GTGCATGTGCGAGGCAATTGCAACAGATCATGTGAGAAACGAGAGCAATAGTTTGACGGGAG	360
301	Db	GTGCATGTGCGAGGCAATTGCAACAGATCATGTGAGAAACGAGAGCAATAGTTTGACGGGAG	360
361	Qy	GCACAGGAGCAACAGTTCAGAGGGAGCTCAGNACTTGCCCTCAACAGTGGCGCCTTAG	420
361	Db	GCACAGGAGCAACAGTTCAGAGGGAGCTCAGNACTTGCCCTCAACAGTGGCGCCTTAG	420
421	Qy	GGCACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCAGAGACAGATCTAAACACC	480
421	Db	GGCACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCAGAGACAGATCTAAACACC	480
481	Qy	TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGCTATTATCTATG	540
481	Db	TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGCTATTATCTATG	540
541	Qy	GTTATGTTTAGTTTTGGTAAATAAAGACATCATCTATATGAATGTTGATCGTGTTA	600
541	Db	GTTATGTTTAGTTTTGGTAAATAAAGACATCATCTATATGAATGTTGATCGTGTTA	600
601	Qy	ACTAAGGCAAGCTTAGTTTATATGACGACCTTTTAGAGTGCTTTTATGGCGCTGCTCTATGT	660
601	Db	ACTAAGGCAAGCTTAGTTTATATGACGACCTTTTAGAGTGCTTTTATGGCGCTGCTCTATGT	660
661	Qy	TTTGTGTCGACAGTTGTAAACCATCTTTGAAATTAATAAAAGATCATGTTTTGTT	717
661	Db	TTTGTGTCGACAGTTGTAAACCATCTTTGAAATTAATAAAAGATCATGTTTTGTT	717

RESULT 2

US-09-191-593-6
 ; Sequence 6, Application US/09191593
 ; Patent No. 6835824
 ; GENERAL INFORMATION:
 ; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
 ; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
 ; APPLICANT: BANNON, Gary A
 ; TITLE OF INVENTION: PEANUT ALLERGENS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 67
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Head, Johnson & Kachigian
 ; STREET: 112 W. Center St., Suite 230
 ; CITY: Fayetteville
 ; STATE: Arkansas AR
 ; COUNTRY: United States of America
 ; ZIP: 72701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: MS-DOS 6.2
 ; SOFTWARE: Wordperfect 6.0C
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/191,593
 ; FILING DATE: 13 NOVEMBER 1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/717,933
 ; FILING DATE: 23 SEPTEMBER 1996
 ; APPLICATION NUMBER: US 07/998,377
 ; FILING DATE: 30 DECEMBER 1992
 ; APPLICATION NUMBER: US 08/158,704
 ; FILING DATE: 29 NOVEMBER 1993
 ; APPLICATION NUMBER: US 60/009,455
 ; FILING DATE: 29 DECEMBER 1995
 ; APPLICATION NUMBER: US 08/610,424
 ; FILING DATE: 04 MARCH 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ALEXANDER, DANIEL R
 ; REGISTRATION NUMBER: 32,604
 ; REFERENCE/DOCKET NUMBER: ARK00895601B
 ; TELECOMMUNICATION INFORMATION:

Qy 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300
Db 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300
Qy 301 GTGATGTGCGAGGCACTTGCACAGATCATGGAGAACCAAGCGATAGGTTGCAGGGGAG 360
Db 301 GTGATGTGCGAGGCACTTGCACAGATCATGGAGAACCAAGCGATAGGTTGCAGGGGAG 360
Qy 361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
Db 361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
Qy 421 GGCACACAGCGTTGCCACTTGGACGTCGAAGTGGCGGCGAGACAGATCACTAAACACC 480
Db 421 GGCACACAGCGTTGCCACTTGGACGTCGAAGTGGCGGCGAGACAGATCACTAAACACC 480
Qy 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Qy 541 GTTATGTTTGTAGTTTGTATATAATAAAGATCATCATATATAGTATATATATATATATATG 600
Db 541 GTTATGTTTGTAGTTTGTATATAATAAAGATCATCATATATAGTATATATATATATATATG 600
Qy 601 ACTAAGCGAAGCTTAGCTTATATAGCAGCACCTTTAGAGTGTCTTTATGCGCTGCTATGT 660
Db 601 ACTAAGCGAAGCTTAGCTTATATAGCAGCACCTTTAGAGTGTCTTTATGCGCTGCTATGT 660
Qy 661 TTTGTTCTGCTGAGAGTTGTAACCACTTTGAAATATATAAAGATCATCATGTTTGT 717
Db 661 TTTGTTCTGCTGAGAGTTGTAACCACTTTGAAATATATAAAGATCATCATGTTTGT 717

RESULT 3

US-09-191-593-9
; Sequence 9, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996

ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: identified as Ara h II cDNA clone
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: No. 6835824 applicable
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE: seed
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE: seed cDNA
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY: florunner seed cDNA expression
LIBRARY: library in Uni-ZAP XR vector
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-9

Query Match 100.0%; Score 717; DB 4; Length 717;
Best Local Similarity 100.0%; Pred No. 2,9e-189;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGCGCATCTCGGAGGCA 60
Db 1 GCTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGCGCATCTCGGAGGCA 60
Qy 61 GCAGTGGGAACCTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAGCTGAG 120
Db 61 GCAGTGGGAACCTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAGCTGAG 120
Qy 121 GCCTCGGAGCAACATCTCATGCAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180

Db	121	GCCTCGCGAGCAACATCTCATGAGAAAGATCCAAACGTGACGAGGATTCATATGAACGGGA	180
Qy	181	CCGTAACAGCCCTAGTCAGGATCCGTACAGCCTAGTCCATATGATCGAGAGCGCGCTGG	240
Db	181	CCGTAACAGCCCTAGTCAGGATCCGTACAGCCTAGTCCATATGATCGAGAGCGCGCTGG	240
Qy	241	ATCCTCTCAGCACCAAGAGAGGTTGTCATATGAGCTGAACGAGTTTGGAGAACCAACAAAG	300
Db	241	ATCCTCTCAGCACCAAGAGAGGTTGTCATATGAGCTGAACGAGTTTGGAGAACCAACAAAG	300
Qy	301	GTGATGTGCGAGGATTCGACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAG	360
Db	301	GTGATGTGCGAGGATTCGACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAG	360
Qy	361	GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG	420
Db	361	GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG	420
Qy	421	GGCACCAACAGCTTGGGACTTGGACGTCGAAAGTGGCGGCGAGACAGATCTAAACACC	480
Db	421	GGCACCAACAGCTTGGGACTTGGACGTCGAAAGTGGCGGCGAGACAGATCTAAACACC	480
Qy	481	TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATCTATG	540
Db	481	TATCTCAAAAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATCTATG	540
Qy	541	GTTATGTTTGGTAAATAAAGATCATCATATGATGATGATGATGATGATGATGATGATG	600
Db	541	GTTATGTTTGGTAAATAAAGATCATCATATGATGATGATGATGATGATGATGATGATG	600
Qy	601	ACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGTCTTATGCGGCTGTCTATGT	660
Db	601	ACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGTCTTATGCGGCTGTCTATGT	660
Qy	661	TTTGTGCTCGAGTTGTAACCATCTTGAATAATAAATAAAGATCATGTTTGT	717
Db	661	TTTGTGCTCGAGTTGTAACCATCTTGAATAATAAATAAAGATCATGTTTGT	717
RESULT 4			
US-09-106-872A-1			
; Sequence 1, Application US/09106872A			
; Patent No. 6486311			
; GENERAL INFORMATION:			
; APPLICANT: Burks Jr., A. Wesley			
; APPLICANT: Stanley, J. Steven			
; APPLICANT: Cockrell, Gael			
; APPLICANT: King, Nina E.			
; APPLICANT: Sampson, Hugh A.			
; APPLICANT: Helm, Ricki M.			
; APPLICANT: Bannon, Gary A.			
; TITLE OF INVENTION: Peanut Allergens and Methods			
; FILE REFERENCE: HS 103 CIP			
; CURRENT APPLICATION NUMBER: US/09/106, 872A			
; CURRENT FILING DATE: 1999-06-29			
; PRIOR APPLICATION NUMBER: PCT/US96/15222			
; PRIOR FILING DATE: 1996-09-23			
; NUMBER OF SEQ ID NOS: 23			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 743			
; TYPE: DNA			
; ORGANISM: Arachis hypogaea			
US-09-106-872A-1			
Query Match 99.6%; Score 713.8; DB 4; Length 743;			
Best Local Similarity 99.7%; Pred. No. 2.3e-188;			
Matches 715; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	GCTCACCATACTAGTAGCCCTCGCCCTTTCCCTCCCTCGCTGCGCCACGACATCTCGGAGGCA	60
Db	6	GCTCACCATACTAGTAGCCCTCGCCCTTTCCCTCCCTCGCTGCGCCACGACATCTCGGAGGCA	65
; ATTORNEY/AGENT INFORMATION:			
; FILING DATE: Concurrently herewith			
; APPLICATION NUMBER: US/08/618,911			
; CURRENT APPLICATION DATA:			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; COMPUTER: IBM PC compatible			
; MEDIUM TYPE: Floppy disk			
; COMPUTER READABLE FORM:			
; ZIP: 50309			
; STATE: Iowa			
; COUNTRY: USA			
; CITY: Des Moines			
; STREET: 700 Capital Square, 400 Locust Street			
; ADDRESS: PIONEER HI-BRED INTERNATIONAL, INC.			
; CORRESPONDENCE ADDRESS:			
; NUMBER OF SEQUENCES: 13			
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN			
; SEEDS			
; APPLICANT: Hu, David			
; APPLICANT: Coughlan, Sean			
; APPLICANT: Hastings, Craig			
; APPLICANT: Jung, Rudolf			
; GENERAL INFORMATION:			
; Patent No. 5850016			
; Sequence 3, Application US/08618911			
US-08-618-911-3			
RESULT 5			
Qy	661	TTTGTGCTCGAGTTGTAACCATCTTGAATAATAAATAAAGATCATGTTTGT	717
Db	666	TTTGTGCTCGAGTTGTAACCATCTTGAATAATAAATAAAGATCATGTTTGT	722


```
QY 489 AAAAAAGAAAAAGAAAAAGAAA 512
Db 1089 RRRRRRRRRRRRRRRRRRRA 1066

RESULT 10
US-08-618-911-1
; Sequence 1, Application US/08618911
; Patent No. 5850016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; TITLE OF INVENTION: SEES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..474
US-08-618-911-1

Query Match 7.6%; Score 54.4; DB 2; Length 723;
Best Local Similarity 61.8%; Pred. No. 4.3e-05;
Matches 123; Conservative 0; Mismatches 71; Indels 6; Gaps 2;

QY 251 CACCAAGAGAGGTGTTCATGAGCTGAACGAGTTTGAGAACACCAAGAGTGCATGTGC 310
Db 277 CACATCGAAGTGTCTGACGCGAATGAGGAGCT---GAAAGCCCCATATGCCAGTGC 333
QY 311 GAGGCATTGCAACAGATCATGTGGAGAACACGAGCGATAGGTTCAGGGGAGGCAACAGGAG 370
Db 334 AAAGCGCTACAGAAGATAATGGATACCAGAGCGGAGCAACTGGAGGGGAAGGAGAAG 393
QY 371 CAACAGTTCAAGGAGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCCCTTAGGGCACCACAG 430
Db 394 ---CAGATGGAGAGAGAGCTCATGAACCTTGGCTATTAGGTGCAGGTGGGACCCATGATA 450
QY 431 CGTTGCGACTTGGACGTCGA 450
Db 451 GGGTGGCACTTGCCTCCGA 470

RESULT 11
US-09-949-016-12026/c
; Sequence 12026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12026
; LENGTH: 168104
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(168104)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12026

Query Match 6.2%; Score 44.4; DB 4; Length 168104;
Best Local Similarity 59.5%; Pred. No. 0.35;
Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 386 GAGCTCAGAACTTGCCTCAACAGTGGCGCTTAGGGCACCACAGGTTGCGACTTGGAC 445
Db 120712 GAGCCAGGAAGTCAAGGCTGCGAGTGAGCCATGATCACACCACTGCAITCCAGCCTGGGT 120653
QY 446 GTCGAAAGTGGCGGAGACAGACAGATATACTAAAAACACCTATCTCAAAAAAGAAAGAAAAGA 505
Db 120652 GACAGATAAGACCCTGTCTCAAAAAACAAACCCCAAAAAA 120593
QY 506 AAAGAA 511
Db 120592 GAAGAA 120587

RESULT 12
US-09-949-016-16554/c
; Sequence 16554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16554
; LENGTH: 168105
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(168105)
; OTHER INFORMATION: n = A,T,C or G
```



```
/ LOCATION: (779676)..(779676)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (855539)..(855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (871619)..(871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1113881)..(1113881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349473)..(1349473)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349491)..(1349491)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1470091)..(1470091)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1569020)..(1569020)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1602912)..(1602912)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1603734)..(1603734)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1637998)..(1637998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1664854)..(1664855)
/ OTHER INFORMATION: n equals a, t, c, or g
/ US-08-916-421B-1
```

```
Query Match 6.2%; Score 44.4; DB 4; Length 1664976;
Best Local Similarity 52.6%; Pred. No. 1.1;
Matches 120; Conservative 0; Mismatches 106; Indels 2; Gaps 1;

QY 488 AAAAAAGAAAAGAAAAGAAAATAGCTTATATATAGCTATTATCTATCGTTATGT 547
    |||||
Db 433470 AAAAAAGAAAAGTGTATTAGTTAGTAATTAATCTACATGGCTTTACATGATGAAA 433529

QY 548 TTAGTTTTGGTAATAATAAAGATCATCATATATGAATGTTGTATCGTGT--AACTAA 605
    |||||
Db 433530 TAAGTTTTTGCAACAAAACAAAATATTGTATGATTTGATTTTCGTATTTCAAAGAA 433589

QY 606 GGCAAGCTTAGGTATATATAGCACCTTTAGAGTGCTTTTATGCGTGTGCTATGTTTGT 665
    |||||
Db 433590 ACATCAATACAGATATTGTAGGTTATTAAAGTAAGCAAGTCATGCTTTTGTAGTAT 433649

QY 666 TGCTGCAGAGTTGTAACCATCTTGAATAATATATAAAAAGATCATGTTT 713
    |||||
Db 433650 AGTTGAACCATTTATACCACCTTCATATAGGAATAAGTAAAGTATGTT 433697
```

```
RESULT 14
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
```


QY 666 TGCTGCAGAGTTGTAACCATCTTTGAAATAATATAAAAAAGATCATGTTT 713
Db 433650 AGTTGAACCATTTATACCACTTTCATATAGGAATAAAGTAAGATGTTT 433697

RESULT 15
US-09-949-016-14026/c
; Sequence 14026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14026
; LENGTH: 56665
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14026

Query Match 6.2%; Score 44.2; DB 4; Length 56665;
Best Local Similarity 56.6%; Pred. No. 0.24;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 385 GGAGCTCAGGAATTGGCTCAACAGTGGCGCTTAGGGCACACACAGGTTGGACTTGA 444
Db 31326 GGAACCCAGGGAGGTGGAGGTTGCAGTGAGCAGAGATCGCACCACTGCCTCCAGCTTGGG 31267
QY 445 CGTCGAAAGTGGCGGAGAGACAGATACTAAACACCTATCTCAAAAAAGAAAAAG 504
Db 31266 CGACAGAGTGAGACTCGTCTCAGAAAAACACACACAAAAAAGGAA 31207
QY 505 AAAAGAAAATAGCTTATATATAAGC 529
Db 31206 GAAAGAAAATAGCTTGTCAATCAAC 31182

Search completed: August 24, 2005, 03:57:02
Job time : 152.962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 433.422 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
 - 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
 - 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
 - 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
 - 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
 - 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
 - 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
 - 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717	100.0	717	10	US-09-731-375A-2
2	717	100.0	717	20	US-10-728-323-2
3	717	100.0	717	20	US-10-728-051-2
4	717	100.0	717	21	US-10-899-551-3
5	474	66.1	474	14	US-10-228-806-3
6	474	66.1	474	17	US-10-100-303A-62
7	462	64.4	682	21	US-10-958-324-3

Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 62, Appli
Sequence 3, Appli

8	462	64.4	1162	21	US-10-958-324-1	Sequence 1, Appli
9	80	11.2	80	21	US-10-958-324-7	Sequence 7, Appli
10	65.6	9.1	1013	18	US-10-424-599-26383	Sequence 26383, A
11	63.4	8.8	770	17	US-10-302-633-1	Sequence 1, Appli
12	62	8.6	62	21	US-10-958-324-8	Sequence 8, Appli
13	54.4	7.6	883	18	US-10-424-599-121530	Sequence 121530,
14	49.8	6.9	62658	19	US-10-322-281-420	Sequence 420, App
15	49.6	6.9	52316	9	US-09-747-810-1	Sequence 1, Appli
16	49.6	6.9	169739	19	US-10-450-826-93	Sequence 93, Appl
17	49	6.8	349	18	US-10-424-599-133074	Sequence 133074,
18	49	6.8	62001	19	US-10-316-459-13	Sequence 13, Appl
19	48.2	6.7	7758	15	US-10-311-455-1076	Sequence 1076, Ap
20	48	6.7	6509	15	US-10-311-455-159	Sequence 159, App
21	47.8	6.7	1237	13	US-10-027-632-122633	Sequence 122633,
22	47.8	6.7	1237	17	US-10-027-632-122633	Sequence 122633,
23	47.6	6.6	451	18	US-10-424-599-24494	Sequence 24494, A
24	47.4	6.6	325	18	US-10-424-599-76496	Sequence 76496, A
25	46	6.4	1069	21	US-10-956-157-107	Sequence 107, App
26	46	6.4	1069	21	US-10-956-157-5342	Sequence 5342, Ap
27	45.6	6.4	6063	15	US-10-240-453-287	Sequence 287, App
28	45.4	6.3	5823	15	US-10-240-453-256	Sequence 256, App
29	45.2	6.3	121724	19	US-10-450-826-50	Sequence 50, Appl
30	45	6.3	48203	21	US-10-741-601-5711	Sequence 5711, Ap
31	45	6.3	48203	21	US-10-741-600-17815	Sequence 17815, A
32	44.4	6.2	6385	14	US-10-239-676-58	Sequence 58, Appl
33	44.4	6.2	6385	15	US-10-240-453-60	Sequence 60, Appl
34	44	6.1	812	13	US-10-027-632-29973	Sequence 29973, A
35	44	6.1	812	17	US-10-027-632-29973	Sequence 29973, A
36	43.8	6.1	7566	10	US-09-764-891-10010	Sequence 10010, A
37	43.8	6.1	174448	13	US-10-087-192-148	Sequence 148, App
38	43.6	6.1	60815	13	US-10-087-192-52	Sequence 52, Appl
39	43.4	6.1	5919	15	US-10-311-455-362	Sequence 362, App
40	43.4	6.1	5919	17	US-10-221-613-64	Sequence 64, Appl
41	43.4	6.1	19696	10	US-09-764-891-9327	Sequence 9327, Ap
42	43.4	6.1	19696	14	US-10-091-572-874	Sequence 874, App
43	43.4	6.1	24173	10	US-09-764-891-9294	Sequence 9294, Ap
44	43.4	6.1	24173	14	US-10-091-572-855	Sequence 855, App
45	43.2	6.0	151	18	US-10-424-599-107242	Sequence 107242,

ALIGNMENTS

RESULT 1
US-09-731-375A-2
; Sequence 2, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731,375A
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-731-375A-2

Query Match	100.0%	Score 717;	DB 10;	Length 717;
Best Local Similarity	100.0%	Pred. No. 2.4e-178;		
Matches 717;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCACGATCTGCGAGGCA	60
DB	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCACGATCTGCGAGGCA	60
QY	61	GCAGTGGGAATCCACAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGCGGCGACCTGAG	120

Db 61 GCAGTGGGAACTCCAAGGACAGAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAG 120
QY 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAAAGCTGACGAGGATTCATATCAACGGGA 180
Db 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAAAGCTGACGAGGATTCATATCAACGGGA 180
QY 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
QY 241 ATCCTCTCAGCACCAAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 241 ATCCTCTCAGCACCAAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
QY 301 GTGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGAGATGTTGACGGGGAG 360
Db 301 GTGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGAGATGTTGACGGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCCTTAG 420
QY 421 GGACACACAGCGTTGGAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATACCTAAACACC 480
Db 421 GGACACACAGCGTTGGAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATACCTAAACACC 480
QY 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATCTATG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATCTATG 540
QY 541 GTTATGTTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
Db 541 GTTATGTTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCCTTTATGCGGCTGCTATGT 660
Db 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCCTTTATGCGGCTGCTATGT 660
QY 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAAATAATATAAAAGATCATGTTTGT 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAAATAATATAAAAGATCATGTTTGT 717

RESULT 2

US-10-728-323-2
; Sequence 2, Application US/10728323
; Publication No. US20040208894A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0232
; CURRENT APPLICATION NUMBER: US/10/728,323
; CURRENT FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea

US-10-728-323-2

Query Match 100.0%; Score 717; DB 20; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.4e-178;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCACCATATAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTCTCGAGGCA 60
Db 1 GCTCACCATATAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGCACTCTCGAGGCA 60
QY 61 GCAGTGGGAACTCCAAGGACAGAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAG 120
Db 61 GCAGTGGGAACTCCAAGGACAGAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAG 120
QY 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAAAGCTGACGAGGATTCATATCAACGGGA 180
Db 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAAAGCTGACGAGGATTCATATCAACGGGA 180
QY 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
QY 241 ATCCTCTCAGCACCAAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 241 ATCCTCTCAGCACCAAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
QY 301 GTGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGAGATGTTGACGGGGAG 360
Db 301 GTGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGAGATGTTGACGGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCCTTAG 420
QY 421 GGACACACAGCGTTGGAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATACCTAAACACC 480
Db 421 GGACACACAGCGTTGGAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATACCTAAACACC 480
QY 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATAGCTTATATCTATG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGTATATATAGCTTATATAGCTTATATCTATG 540
QY 541 GTTATGTTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
Db 541 GTTATGTTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCCTTTATGCGGCTGCTATGT 660
Db 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCCTTTATGCGGCTGCTATGT 660
QY 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAAATAATATAAAAGATCATGTTTGT 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAAATAATATAAAAGATCATGTTTGT 717

RESULT 3

US-10-728-051-2
; Sequence 2, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10/728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea

US-10-728-051-2

Query Match	100.0%	Score	717;	DB	20;	Length	717;
Best Local Similarity	100.0%	Pred.	No. 2.4e-178;				
Matches	717;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCCCTCGCTCGCCGCCACCGATCTCGGAGCA	60				
Db	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCCCTCGCTCGCCGCCACCGATCTCGGAGCA	60				
Qy	61	GCAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGCAACCTGAG	120				
Db	61	GCAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGCAACCTGAG	120				
Qy	121	GCCCTCGGAGCAATCTCATCGCAGAGATGCCAAAGTGAACGAGATTTCATATGAACGGGA	180				
Db	121	GCCCTCGGAGCAATCTCATCGCAGAGATGCCAAAGTGAACGAGATTTCATATGAACGGGA	180				
Qy	181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCTCGGAGAGCGCTGG	240				
Db	181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCTCGGAGAGCGCTGG	240				
Qy	241	ATCCTCTCAGCAACCAAGAGAGGTGTTCGAATGAGCTGAAACGAGTTTCAGAAACCAACAAAG	300				
Db	241	ATCCTCTCAGCAACCAAGAGAGGTGTTCGAATGAGCTGAAACGAGTTTCAGAAACCAACAAAG	300				
Qy	301	GTGCATGTGGAGGCATTTGCAACAGATCATGAGAGAAACAGAGCGATAGTTTGCAGGGGAG	360				
Db	301	GTGCATGTGGAGGCATTTGCAACAGATCATGAGAGAAACAGAGCGATAGTTTGCAGGGGAG	360				
Qy	361	GCACAGGAGCACAAGTTCAGAGGGGAGCTCAGGAACTTGCCCTCAAACAGTCCGCGCCCTAG	420				
Db	361	GCACAGGAGCACAAGTTCAGAGGGGAGCTCAGGAACTTGCCCTCAAACAGTCCGCGCCCTAG	420				
Qy	421	GGCACCAACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGACAGACAGATACATAAACACC	480				
Db	421	GGCACCAACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGACAGACAGATACATAAACACC	480				
Qy	481	TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	540				
Db	481	TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	540				
Qy	541	GTTATGTTTGTGTTTGTGTAATAATAAGATCATCACTATATGAATGTTGTGATCGTGTTA	600				
Db	541	GTTATGTTTGTGTTTGTGTAATAATAAGATCATCACTATATGAATGTTGTGATCGTGTTA	600				
Qy	601	ACTAAGCAACGCTTAGGTATATCAGCACCTTTAGAGTGCCTTTTATGCGCTTGTCTATGT	660				
Db	601	ACTAAGCAACGCTTAGGTATATCAGCACCTTTAGAGTGCCTTTTATGCGCTTGTCTATGT	660				
Qy	661	TTTGTGTCGCAGAGTTGTGAACCACTTGTGAATAATAATAAAGATCATGTTTTTGT	717				
Db	661	TTTGTGTCGCAGAGTTGTGAACCACTTGTGAATAATAATAAAGATCATGTTTTTGT	717				

RESULT 4

```

US/10-899-551-3
/ Sequence 3, Application US/10899551
/ Publication No. US20050063994A1
/ GENERAL INFORMATION:
/ APPLICANT: Caplan, Michael J.
/ APPLICANT: Burke, A. Wesley
/ APPLICANT: Sampson, Hugh A.
/ APPLICANT: Howard, Sosin B.
/ APPLICANT: Bottomly, Kim H.
/ TITLE OF INVENTION: Methods and Re
/ FILE REFERENCE: 2002834-0233
/ CURRENT APPLICATION NUMBER: US/10/
/ CURRENT FILING DATE: 2004-07-26
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: Patent version 3.2
/ SEQ ID NO 3
/ LENGTH: 717

```

; TYPE: DNA
; ORGANISM: species Arachis hypogaea
US-10-899-551-3

Query Match	100.0.0%;	Score	717;	DB	21;	Length	717;
Best Local Similarity	100.0.0%;	Prod.	No. 2.4e-178;				
Matches	717;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GCTCACCATACTAGTAGCCCTCGGCCCTTTTCCTCTCGCTGCCACGATCTCGAGGCA	60				
Db	1	GCTCACCATACTAGTAGCCCTCGGCCCTTTTCCTCTCGCTGCCACGATCTCGAGGCA	60				
Qy	61	GCAGTGGGAACCTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG	120				
Db	61	GCAGTGGGAACCTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG	120				
Qy	121	GCCCTGGGAGCAATCTCATGCAAGAAGATCCAAAGTGCAGAGGATTCATATGAACGGGA	180				
Db	121	GCCCTGGGAGCAATCTCATGCAAGAAGATCCAAAGTGCAGAGGATTCATATGAACGGGA	180				
Qy	181	CCGTCACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGG	240				
Db	181	CCGTCACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGG	240				
Qy	241	ATCCTCTCAGCACCAAGAGAGGTGTTCATATGAGCTGAACGAGTTTGAGAAACCAACG	300				
Db	241	ATCCTCTCAGCACCAAGAGAGGTGTTCATATGAGCTGAACGAGTTTGAGAAACCAACG	300				
Qy	301	GTGCATGTGCGAGGCATTGCAACAGATCATGAGAAACAGAGCGATAGTTGACAGGGAG	360				
Db	301	GTGCATGTGCGAGGCATTGCAACAGATCATGAGAAACAGAGCGATAGTTGACAGGGAG	360				
Qy	361	GCACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCTCAACAGTGGCGCCTTAG	420				
Db	361	GCACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCTCAACAGTGGCGCCTTAG	420				
Qy	421	GGCACACACAGCTTGCAGCTTGGACGTCGAAAGTGGCGGACAGACAGATACATAAACCC	480				
Db	421	GGCACACACAGCTTGCAGCTTGGACGTCGAAAGTGGCGGACAGACAGATACATAAACCC	480				
Qy	481	TATCTCAAAAAAGAAAAAGAAAAATAGCTTATATAAAGCTATTATCTATG	540				
Db	481	TATCTCAAAAAAGAAAAAGAAAAATAGCTTATATAAAGCTATTATCTATG	540				
Qy	541	GTTATGTTTGTGTTTGTGTAATAAAGATCATCACTATATGAATGTGTGATCGTGTTA	600				
Db	541	GTTATGTTTGTGTTTGTGTAATAAAGATCATCACTATATGAATGTGTGATCGTGTTA	600				
Qy	601	ACTAAGGCAAGCTTAGGTATATGAGCACCTTTAGAGTGCCTTTTATGGCGTTTGTCTATGT	660				
Db	601	ACTAAGGCAAGCTTAGGTATATGAGCACCTTTAGAGTGCCTTTTATGGCGTTTGTCTATGT	660				
Qy	661	TTTGTGTGTCAGAGTTGTAAACCATCTTGAAATAATAAAGATCATGTTTTGTT	717				
Db	661	TTTGTGTGTCAGAGTTGTAAACCATCTTGAAATAATAAAGATCATGTTTTGTT	717				

RESIT.T 5

```

0001 5
0002 US-10-228-806-3
0003 ; Sequence 3, Application US/10228806
0004 ; Publication No. US20030049237A1
0005 ; GENERAL INFORMATION:
0006 ; APPLICANT: Bannan, et al.
0007 ; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
0008 ; TITLE OF INVENTION: to Allergy
0009 ; FILE REFERENCE: 2002834-0043
0010 ; CURRENT APPLICATION NUMBER: US/10/228,806
0011 ; CURRENT FILING DATE: 2002-08-26
0012 ; NUMBER OF SEQ ID NOS: 81
0013 ; SOFTWARE: PatentIn Ver. 2.1
0014 ; SEQ ID NO 3
0015 ; LENGTH: 474
0016 ; TYPE: DNA

```

; ORGANISM: Arachis hypogaea		; ORGANISM: Arachis hypogaea	
US-10-228-806-3		US-10-958-324-3	
Query Match 66.1%; Score 474; DB 14; Length 474;		Query Match 64.4%; Score 462; DB 21; Length 682;	
Best Local Similarity 100.0%; Pred. No. 2e-114;		Best Local Similarity 100.0%; Pred. No. 3.6e-111;	
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 2	CTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 61	QY 1	GCTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60
DB 1	CTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60	DB 67	GCTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 126
QY 62	CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAGG 121	QY 61	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAG 120
DB 61	CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAGG 120	DB 127	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAG 186
QY 122	CCCTGGGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 181	QY 121	GCCTCGCGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 180
DB 121	CCCTGGGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 180	DB 187	GCCTCGCGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 246
QY 182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241	QY 181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 240
DB 181	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 240	DB 247	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 306
QY 302	TGATGTGCGAGGCAATTCGAATGAGCTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 301	QY 241	ATCCTCTCAGCAGCAAGAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300
DB 301	TGATGTGCGAGGCAATTCGAATGAGCTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300	DB 307	ATCCTCTCAGCAGCAAGAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 366
QY 362	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGCAGCTTAGG 421	RESULT 6	
DB 361	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGCAGCTTAGG 420	US-10-100-303A-62	
QY 422	GCACCACAGCGTTGCGACTTGGACGCTCGAAAGTGGCGGCGAGACAGATACTAA 475	; Sequence 62, Application US/10100303A	
DB 421	GCACCACAGCGTTGCGACTTGGACGCTCGAAAGTGGCGGCGAGACAGATACTAA 474	; Publication No. US20030202980A1	
RESULT 7		; GENERAL INFORMATION:	
US-10-958-324-3		; APPLICANT: Caplan, et al.	
; Sequence 3, Application US/10958324		; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction	
; Publication No. US20050114924A1		; FILE REFERENCE: 2002834-0166	
; GENERAL INFORMATION:		; CURRENT APPLICATION NUMBER: US/10/100,303A	
; APPLICANT: DODO, HORTENSE W.		; NUMBER OF SEQ ID NOS: 138	
; APPLICANT: ARNTZEN, CHARLES J.		; SOFTWARE: PatentIn Ver. 2.1	
; APPLICANT: KONAN, KOFFI N'DA		; SEQ ID NO 62	
; APPLICANT: VIOQUEZ, OLGA		; LENGTH: 474	
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN		; TYPE: DNA	
; FILE REFERENCE: 072121/0104		; ORGANISM: Arachis hypogaea	
; CURRENT APPLICATION NUMBER: US/10/958,324		US-10-958-324-3	
; CURRENT FILING DATE: 2004-10-06		Query Match 66.1%; Score 474; DB 17; Length 474;	
; PRIOR APPLICATION NUMBER: US/09/715,036		Best Local Similarity 100.0%; Pred. No. 2e-114;	
; PRIOR FILING DATE: 2000-11-20		Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
; PRIOR APPLICATION NUMBER: 60/167,255		QY 2	CTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 61
; PRIOR FILING DATE: 1999-11-19		DB 1	CTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60
; NUMBER OF SEQ ID NOS: 8		QY 62	CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAGG 121
; SOFTWARE: PatentIn Ver. 2.1		DB 61	CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAGG 120
; SEQ ID NO 3		QY 122	CCCTGGGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 181
; LENGTH: 682		DB 121	CCCTGGGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 180
; TYPE: DNA		QY 182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
; ORGANISM: Arachis hypogaea		DB 181	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 240
US-10-958-324-3		QY 242	TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 301
Query Match 64.4%; Score 462; DB 21; Length 682;		DB 241	TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300
Best Local Similarity 100.0%; Pred. No. 3.6e-111;		QY 302	TGATGTGCGAGGCAATTCGAATGAGCTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 361
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		DB 301	TGATGTGCGAGGCAATTCGAATGAGCTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 360
QY 1	GCTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60	QY 362	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGCAGCTTAGG 421
DB 67	GCTCACCATACTAGTAGCCCTTCGCCCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 126	DB 361	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGCAGCTTAGG 420
QY 61	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAG 120	QY 422	GCACCACAGCGTTGCGACTTGGACGCTCGAAAGTGGCGGCGAGACAGATACTAA 475
DB 127	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTCTGAG 186	DB 421	GCACCACAGCGTTGCGACTTGGACGCTCGAAAGTGGCGGCGAGACAGATACTAA 474
QY 121	GCCTCGCGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 180	RESULT 6	
DB 187	GCCTCGCGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGATTCATATGAACGGGAC 246	US-10-100-303A-62	
QY 181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 240	; Sequence 62, Application US/10100303A	
DB 247	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 306	; Publication No. US20030202980A1	
QY 241	ATCCTCTCAGCAGCAAGAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 300	; GENERAL INFORMATION:	
DB 307	ATCCTCTCAGCAGCAAGAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAAAG 366	; APPLICANT: Caplan, et al.	

Qy 301 GTGCATGTGGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGGTTGCGAGGGAG 360
Db 367 GTGCATGTGGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGGTTGCGAGGGAG 426
Qy 361 GCAACAGAGCAACAGATTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 420
Db 427 GCAACAGAGCAACAGATTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 486
Qy 421 GGACACACAGGTTGCGACTTGGACGTCGAAAGTGGGGCGAG 462
Db 487 GGACACACAGGTTGCGACTTGGACGTCGAAAGTGGGGCGAG 528

RESULT 8
US-10-958-324-1
; Sequence 1, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (110)..(730)
US-10-958-324-1

Query Match 64.4%; Score 462; DB 21; Length 1162;
Best Local Similarity 100.0%; Pred. No. 4.7e-111;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCAATCTGCGAGGCA 60
Db 118 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCAATCTGCGAGGCA 177
Qy 61 GCAGTGGGAATCTCAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGGAACCTTAG 120
Db 178 GCAGTGGGAATCTCAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGGAACCTTAG 237
Qy 121 GCCTGCGAGCAATCTCATGCGAGAGATCCACGTCGAGGAGGATTCATAGACGGGA 180
Db 238 GCCTGCGAGCAATCTCATGCGAGAGATCCACGTCGAGGAGGATTCATAGACGGGA 297
Qy 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 240
Db 298 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 357
Qy 241 ATCTCTTCAGACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
Db 358 ATCTCTTCAGACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 417
Qy 301 GTGCATGTGGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGGTTGCGAGGGAG 360
Db 418 GTGCATGTGGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGGTTGCGAGGGAG 477
Qy 361 GCAACAGAGCAACAGATTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 420
Db 478 GCAACAGAGCAACAGATTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 537

Qy 421 GGACACACAGGTTGCGACTTGGACGTCGAAAGTGGGGCGAG 462
Db 538 GGACACACAGGTTGCGACTTGGACGTCGAAAGTGGGGCGAG 579

RESULT 9
US-10-958-324-7
; Sequence 7, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-958-324-7

Query Match 11.2%; Score 80; DB 21; Length 80;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCACGCAATCTCGAGGCGAGCGAGGAA 70
Db 1 CTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCACGCAATCTCGAGGCGAGCGAGGAA 60

Qy 71 CTCACAGGAGACAGAGATG 90
Db 61 CTCACAGGAGACAGAGATG 80

RESULT 10
US-10-424-599-26383
; Sequence 26383, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 26383
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1013)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123825C.1
US-10-424-599-26383

Db 611 GTGAATGGTATGCTGATGACCTACTAGTAGCAATGGAAGCACTTAGAGTGCT 670
Qy 642 TT---TATGGCGTTGTCTATGTTTGTGTCGACAGAGTTGTAACCATCTTTGAAAT 693
Db 671 TTGTGCAATGGCTTGCCTCTGTTTGTGAGACTTTTGTAAATGTTTTCGAGTTTAAAT 725

RESULT 12

US-10-958-324-8
; Sequence 8, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIQUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 62
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-958-324-8

Query Match 8.6%; Score 62; DB 21; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 301 GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACAGAGCGATAGTTGACGGGGAG 360
Db 1 GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACAGAGCGATAGTTGACGGGGAG 60
Qy 361 GC 362
Db 61 GC 62

RESULT 13

US-10-424-599-121530
; Sequence 121530, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121530
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80749C.1
US-10-424-599-121530

Query Match 7.6%; Score 54.4; DB 18; Length 883;

Best Local Similarity 61.5%; Pred. No. 0.001;
Matches 123; Conservative 0; Mismatches 71; Indels 6; Gaps 2;
Qy 251 CACCAAGAGAGAGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGGTCATGTC 310
Db 437 CACATGCAGAAGTGTGCGAGCAATGAGCGAGCT---GAAAGCCCATATGCCAGTGC 493
Qy 311 GAGGCATTTGCAACAGATCATGGAGAACAGAGGATAGTTGAGGGGAGGCAACAGAG 370
Db 494 AAAGCGCTACAGAAGATAATATGATAACAGAGCGAGCAACTGGAGGGGAGGAGAAAG 553
Qy 371 CAACAGTTCAAGAGGAGGAGCTCAGAACTTGCCCTCAACAGTGGCGCTTTAGGGACACAG 430
Db 554 ---CAGATGAGAGAGAGCTCATGAACCTTGCTATTAGGTGACAGTTGGGACCATGATA 610
Qy 431 CGTTGCGACTTGGACGTCGA 450
Db 611 GGGTGGGACTTGTCTCCGA 630

RESULT 14

US-10-322-281-420/c
; Sequence 420, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 62658
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-420

Query Match 6.9%; Score 49.8; DB 19; Length 62658;
Best Local Similarity 54.1%; Pred. No. 0.14;
Matches 125; Conservative 0; Mismatches 102; Indels 4; Gaps 1;
Qy 386 GAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTAGGGCCACACAGCGTTGCGACTTGGAC 445
Db 61081 GAGCCCAAGGAGTTGAGGCTGCGAGTGCAGTGCAGCTGCTGCTCCAGCCTGGGC 61022
Qy 446 GTCGAAAGTGGCGGAGAGACAGATATAACACCTATCTCAAAAAAGAAAGAAAGA 505
Db 61021 GACAGAGCAAGACTCTGTCTCAAAAAAGAAAGAAAGAAAGAAAGA 60962

Qy 506 AAAGAAATAGCTTATATATATAGCTATTATCTATGTTATGTTTGTGTAATAATA 565
Db 60961 AAAGAAAGGA---AAATAGCTATGCTTGTGTTATGGGAGTCAAGGACACACA 60906
Qy 566 AAGATCATCATATATGAATGTGTTGATCGTTAACTAAGGCAAGCTTAG 616
Db 60905 GTGGCAAGGTAAGTTCAATCTTCAGTTAGTTTGGCAAGTTTAAGTTTG 60855

RESULT 15

US-09-747-810-1/c
; Sequence 1, Application US/09747810
; Patent No. US20020012903A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Hideyuki
; APPLICANT: Goldman, Steven A.
; TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGENITOR CELLS
; FILE REFERENCE: 19603/3580
; CURRENT APPLICATION NUMBER: US/09/747,810
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,003

```
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 52216
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-810-1

Query Match      6.9%; Score 49.6; DB 9; Length 52216;
Best Local Similarity 61.7%; Pred.No. 0.15; 49; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 386 GAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTAGGGGCACACAGCGTTGCCACTTGGAC 445
Db 12026 GAACTCAGGAGCTGGAGTTTGCACTGAGCGAGATCGCACCACCTGCATTCCAGCCTGGGC 11967

QY 446 GTCGAAAGTGGCGGCAGAGACAGATACATAACACCTATCTCAAAAAAGAAAAAGAAAGA 505
Db 11966 AACAGAGTGAGACTCTGTCTCAAAAAATAAAAAAAATTTAAAAAAAGAAAAAGAAAGA 11907

QY 506 AAAGAAAA 513
Db 11906 AAAAGAA 11899
```

Search completed: August 24, 2005, 09:47:13
Job time : 435.422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 40.6071 Seconds
(without alignments)
13658.075 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 getcaccatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2103692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlh
-Q=/cgn2.1/USPTO.spool/US10728323/runat.23082005.124355.29200/app.query.fasta.1.4757
-DB=A.Geneseq.16Dec04 -OPMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN.1.1.224 @runat.23082005.124355.29200 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOPT=10 -XGAPEXT=0.5 -FGAPOPT=6
-FGAPEXT=7 -YGAPOPT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq.16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	66.6	157	2 AAW24153	Aaw24153 Peanut al
2	842	66.6	157	2 AAW24164	Aaw24164 Peanut al
3	842	66.6	157	2 AAY15245	Aay15245 Peanut al
4	842	66.6	157	4 AAU04707	Aau04707 Anaphylac
5	842	66.6	157	6 ABUS2464	Abus2464 Peanut Ar
6	842	66.6	157	7 ADG27518	Adg27518 Peanut al
7	838.5	66.3	207	4 AAB82383	Aab82383 Peanut al
8	830	65.6	156	2 AAY40973	Aay40973 Ara h 2 p
9	819	64.7	156	8 ADO38314	Ado38314 Ara h 2 M
10	814	64.3	157	6 ABUS2576	Abus2576 Peanut Ar

11	812	64.2	156	8 ADM12096	Adm12096 Arachis h
12	812	64.2	157	6 ABUS2575	Abus2575 Peanut Ar
13	774	61.2	157	6 ABUS2577	Abus2577 Peanut Ar
14	771	60.9	166	2 AAY40968	Aay40968 Recombina
15	771	60.9	166	6 ABUS2482	Abus2482 Peanut Ar
16	771	60.9	166	7 ADG27536	Adg27536 T7/His-ta
17	699	55.3	166	3 AAB33600	Aab33600 Modified
18	699	55.3	166	4 AAU05035	Aau05035 Modified
19	688.5	54.4	167	4 AAU04710	Aau04710 Modified
20	243	19.2	158	2 AAW23419	Aaw23419 Soybean a
21	243	19.2	158	2 AAY05723	Aay05723 Soybean G
22	243	19.2	158	7 ADH89269	Adh89269 G. max 2S
23	243	19.2	158	8 ADG44004	Adg44004 G. max 2S
24	241.5	19.1	155	2 AAW23418	Aaw23418 Soybean a
25	241.5	19.1	155	7 ADH89267	Adh89267 G. max na
26	241.5	19.1	155	8 ADG44002	Adg44002 G. max 2S
27	229	18.1	158	2 AAW23420	Aaw23420 Chimeric
28	184.5	14.6	323	7 ADH89277	Adh89277 Sunflower
29	184.5	14.6	323	8 ADG44012	Adg44012 H. annuus
30	177.5	14.0	168	6 AAE36076	Aae36076 Flax Conl
31	156	12.3	28	2 AAW24194	Aaw24194 Peanut al
32	156	12.3	158	2 AAW23586	Aaw23586 Mabinlin
33	151	11.9	158	2 ADH89275	Adh89275 Sunflower
34	151	11.9	295	7 ADH89275	Adh89275 Sunflower
35	151	11.9	295	8 ADG44010	Adg44010 H. annuus
36	149.5	11.8	167	8 ADS16393	Ads16393 Grape 2S
37	146.5	11.6	164	1 AAP91892	Aap91892 1kb fragm
38	146.5	11.6	164	1 AAP96144	Aap96144 Sequence
39	146.5	11.6	164	7 ADH89227	Adh89227 A. thalia
40	146.5	11.6	164	8 ADG43962	Adg43962 A. thalia
41	144	11.4	169	6 AAE36077	Aae36077 Flax Conl
42	140.5	11.1	146	6 AAO27313	Aao27313 Brazil nu
43	140.5	11.1	146	8 ADN37294	Adn37294 Brazil nu
44	139	11.0	140	1 AAP91891	Aap91891 Brazil nu
45	139	11.0	155	2 AAW23587	Aaw23587 Mabinlin

ALIGNMENTS

RESULT 1

AAW24153

ID AAW24153 standard; protein; 157 AA.

AC AAW24153;

DT 17-OCT-2003 (revised)

DT 29-DEC-1997 (first entry)

DE Peanut allergen Ara hII.

DE Peanut allergen Ara hII.

KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;

KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;

KW ELISA; analysis; Ara hII.

OS Arachis hypogaea; strain Florunner.

PN WO9724139-A1.

PD 10-JUL-1997.

PF 23-SEP-1996; 96WO-US015222.

PR 29-DEC-1995; 95US-0009455P.

PR 04-MAR-1996; 96US-00610424.

(UYAR-) UNIV ARKANSAS.

PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

WPI; 1997-363453/33.

DR N-PSDB; AAT76615.

PT Peanut allergens Ara hI and Ara hII - used for vaccination and in two-

Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CAACAGAGCAACAGTTCCAGAGGAGCTCAGGAACCTTGCTCAACAGTGCGGCTTAGG 421
 Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCACAGGCTTCGACCTGGAGCTCGAAGTGCGCGCAGACAGACATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 3
 AAY15245
 ID AAY15245 standard; protein; 157 AA.
 AC AAY15245;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-NOV-1999 (first entry)
 XX
 DE Peanut allergen, Ara h 2, amino acid sequence.
 XX
 KW allergy; immune response; transgenic; allergen; epitope;
 KW immunoglobulin E; Ig E; binding site; peanut.
 XX
 OS Arachis hypogaea.
 XX
 PN WO9938978-A1.
 XX
 PD 05-AUG-1999.
 XX
 PF 29-JAN-1999; 99WO-US002031.
 XX
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 27-AUG-1998; 98US-00141220.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 PA (UYNV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) SOSIN H.
 XX
 PI Sosin H, Bannon GA, Burks AW, Sampson HA;
 XX
 DR WPI; 1998-479189/40.
 DR N-PSDB; AA206383.
 XX
 PT Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 XX
 PS Disclosure; Page 38; 46pp; English.
 XX
 CC This is the amino acid sequence of the Ara h 2 protein from Arachis
 CC hypogaea. The Ara h 2 protein has 10 IgE (Immunoglobulin E) binding
 CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
 CC AAY15276). By modifying the IgE binding sites the ability of the allergen
 CC to provoke an immune response is downregulated. The epitopes of the IgE
 CC binding sites can therefore be modified in genetically engineered plants
 CC and animals to elicit less of an allergic response. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 157 AA;

Alignment Scores:
 Pred. No.: 8.2e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 66.56% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAY15245 (1-157)

Qy 2 CTCACACTAGTAGCCCTCGCCCTTTTCCTCCTCGCTGCCACCGCATCTCGAGGCAG 61
 Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 Qy 62 CAGTGGAACTCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCCGAGAGGGCGAACTCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 Qy 122 CCTCGAGCAACATCTCATGCAAGATCCAACTGACCGTACCGAGGATTCATATGAACTGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTTGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrSerProTyrSerProTyr 80
 Qy 242 TCCTCTCAGCAACCAAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
 Qy 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCCAGAGCCATAGTTGCGAGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGGCTTAGG 421
 Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGACATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 4
 AAU04707
 ID AAU04707 standard; protein; 157 AA.
 XX
 AC AAU04707;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Anaphylactic antigen Ara h 2.
 XX
 KW Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse.
 XX
 OS Mus sp.
 XX
 PN WO200140264-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 06-DEC-2000; 2000WO-US033124.
 XX
 PR 06-DEC-1999; 99US-00455294.
 PR 23-JUN-2000; 2000US-021376SP.
 PR 27-SEP-2000; 2000US-0235797P.
 XX
 PA (PANAR-) PANACEA PHARM LLC.
 PA (UYAR-) UNIV ARKANSAS.
 PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
 XX
 PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
 XX
 DR WPI; 2001-381378/40.
 XX
 PT Antigenic fragments useful for reducing anaphylactic risk and reducing
 PT the severity and/or number of allergic symptoms in individuals sensitive
 PT to antigens, have reduced ability to bind Immunoglobulin E.
 XX
 PS Claim 7; Fig 10; 100pp; English.
 XX
 CC The sequence represents the amino acid sequence of anaphylactic antigen
 CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design

CC antigenic peptides having a reduced ability to bind immunoglobulin E
CC (IgE) as compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8.2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x AAU04707 (1-157)

QY 2 CTCACCATAGTAGGCGCTTCCTCTCTGCTGCCACGCGATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAAGAGACAGAAGATGCCAGGCGATCCAGAGGGCGCAACTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArGysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGGAGCAACATCTCATGCGAGAGATCCACGTCGACGAGGATTCATATGAACGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATAGCTGAACAGTTTGAGAACACCAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCGATTCGCAACAGATCATGAGAAACAGAGCGATAGTTTCAGGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTTCAAGAGGAGGAGCTCAGGAACCTTGCTCAACAGTGCAGCTTAGG 421
DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnCysGlyLeuArg 140
QY 422 GCACCAAGCGTTGCGAGCTTGAGCGTGAAGTGGCGGCGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 5
ABU52464
ID ABU52464 standard; protein; 157 AA.
XX
AC ABU52464;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h2 protein.
XX
KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US0009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon CA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI: 2003-018765/01.
DR N-PSDB; ABX70606.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Claim 27; Fig 41; 300pp; English.
XX

CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen (e.g. Ara h1, h2 or h3)

SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8.2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABU52464 (1-157)

QY 2 CTCACCATAGTAGGCGCTTCCTCTCTGCTGCCACGCGATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAAGAGACAGAAGATGCCAGGCGAGCTCGAGAGGCGCAACTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArGysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80

QY 242 TCCTCTCAGCACCAGAGAGGTGTCATGAGCTGAACGAGTTTGAGAACACCAAGG 301
 Db |||||
 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 QY 302 TGCATGTGCGAGGCAATTCACACAGATCATGAGAACACGAGCGATAGTTTCAGGGGAGG 361
 Db |||||
 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
 Db |||||
 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACACAGCGTTCCGACTTGAGCGTGAAGTGGCGGCAGACAGATAC 472
 Db |||||
 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 6
 ADG27518
 ID ADG27518 standard; protein; 157 AA.
 AC
 XX ADG27518;
 DT 26-FEB-2004 (first entry)
 XX
 DE Peanut allergen Ara h2.
 XX
 KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycine A2B1a; Jug n1;
 KW antiallergic; vulnerary; anaphylactic food allergen; IGE; allergy; wound.
 XX
 OS *Arachis hypogaea*.
 XX
 FN US2003202980-A1.
 XX
 PD 30-OCT-2003.
 XX
 PF 18-MAR-2002; 2002US-00100303.
 XX
 PR 29-DEC-1995; 95US-0009455P.
 PR 23-SEP-1996; 96US-0071793P.
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 29-JUN-1998; 98US-00106872.
 PR 27-AUG-1998; 98US-00141220.
 PR 13-NOV-1998; 98US-00191593.
 PR 29-JAN-1999; 99US-00240557.
 PR 29-JAN-1999; 99US-00241101.
 PR 11-FEB-1999; 99US-00248673.
 PR 11-FEB-1999; 99US-00248674.
 PR 02-MAR-1999; 99US-0122450P.
 PR 02-MAR-1999; 99US-0122452P.
 PR 02-MAR-1999; 99US-0122560P.
 PR 02-MAR-1999; 99US-0122565P.
 PR 02-MAR-1999; 99US-0122566P.
 PR 11-MAR-1999; 99US-00267719.
 PR 28-JAN-2000; 2000US-00494096.
 PR 16-MAR-2001; 2001US-0276822P.

(CAPL/) CAPLAN M. J.
 (SOSI/) SOSIN H. B.
 (SAMP/) SAMPSON H.
 (BANN/) BANNON G. A.
 (BURK/) BURKS A. W.
 (COCK/) COCKRELL G.
 (COMP/) CONNAPDRE C. M.
 (CONN/) CONNAUGHTON C.
 (HELM/) HELM R. M.
 (KING/) KING N. E.
 (KOPP/) KOPPER R. A.
 (MALE/) MALEKI S. J.
 (RABU/) RABJOHN P. A.

PA (SHIN/) SHIN D. S.
 PA (STAN/) STANLEY J. S.
 XX
 PI Caplan MJ, Sosin HB, Sampson H, Bannan GA, Burks AW, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabjohn PA, Shin DS, Stanley JS;
 XX
 DR WPI; 2003-875632/81.
 DR N-PSDB; ADG27517.
 XX
 XX New modified anaphylactic food allergen comprising a cysteine residue
 PT which has been modified so that it cannot participate in the disulfide
 PT bond, useful for treating allergic reactions or wounds.
 XX
 PS Claim 27; SEQ ID NO 63; 194pp; English.
 XX
 CC The invention relates to a modified anaphylactic food allergen whose
 CC amino acid sequence is substantially identical to that of a natural
 CC anaphylactic food allergen. The natural anaphylactic food allergen
 CC includes at least one cysteine residue that participates in a disulfide
 CC bond when the natural anaphylactic food allergen is in its native
 CC conformation, except that the cysteine residue has been modified so that
 CC it cannot participate in the disulfide bond. Also included are a method
 CC of making a modified anaphylactic food allergen, a nucleotide molecule
 CC encoding a modified anaphylactic food allergen defined above, a
 CC nucleotide molecule for causing a site specific mutation in a gene
 CC encoding a natural anaphylactic food allergen, a transgenic plant or
 CC animal expressing a modified anaphylactic food allergen defined above, a
 CC method of treating an individual by reducing the clinical response to a
 CC natural anaphylactic food allergen by administering a modified
 CC anaphylactic food allergen and an isolated fragment of peanut allergen
 CC comprising at least 10 consecutive amino acids of ADG27464 or
 CC ADG27465. About 10-17% of the amino acids have been modified in at least
 CC one IGE epitope or all the IGE epitopes recognised when the natural
 CC anaphylactic food allergen is contacted with serum IgE from individual(s)
 CC allergic to the natural anaphylactic food allergen. The invention
 CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
 CC cDNAs), Soybean Glycinin A2B1a and IGE-binding epitopes of the English
 CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
 CC used for treating allergic reactions or wounds. The present sequence
 CC represents a Peanut allergen of the invention (or its fragment).
 XX
 SQ Sequence 157 AA;
 Alignment Scores:
 Pred. No.: 8,2e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 66.56% Indels: 0
 DB: 7 Gaps: 0

US-10-728-323-2 (1-717) x ADG27518 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTCTCTCTCGCTGCCCGCATCTGGAGGAG 61
 Db |||||
 1 LeuthrileuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAAGTCCCAAGGAGACAGAGATGCCAGAGCAGCTCGAGAGGGGGAACCTGAGG 121
 Db |||||
 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCCTGCGAGCAACATCTCATGCGAGAGATCCCAAGTGCAGAGGATTCATATGAAACGGAC 181
 Db |||||
 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CGGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATCATCGAGAGGCGGTGGA 241
 Db |||||
 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
 QY 242 TCCTCTCAGCAACCAAGAGAGGTGTCGAATGACCTGAACGAGTTTCAGAACCAACCAAGG 301
 Db |||||
 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100

XX PI Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
 XX PI Sampson H, Compadre CW, Huang SK, Maleki SJ, Kopper RA;
 XX DR WPI; 1999-551218/46.
 XX PT Tertiary structure of peanut allergen Ara h 1 for protection of a host
 XX PT animal from allergic reaction.
 XX PS Disclosure; Page 104; 193pp; English.
 XX CC The invention provides a tertiary structure for the peanut allergen Ara H
 CC 1. The Ara H 1 allergen is found to contain 23 linear IGE-binding
 CC epitopes. The invention also provides an isolated recombinant peanut
 CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
 CC allergen Ara h 3. Molecules of the invention are used to protect a host
 CC animal from allergic reaction, particularly using a modified allergen
 CC which is less reactive with IGE. The invention may also be used to ensure
 CC that the allergen is not introduced into genetically modified food. The
 CC present sequence represents an Ara h 2 protein fragment. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX SQ Sequence 156 AA;
 Alignment Scores:
 Pred. No.: 2.13e-88 Length: 156
 Score: 830.00 Matches: 155
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 65.61% Indels: 0
 DB: 2 Gaps: 0
 US-10-728-323-2 (1-717) x AAY40973 (1-156)
 QY 2 CTCACCTACTAGTAGCCCTCGCCCTTTCTCTCTCTGCGCCAGCATCTCGAGGAG 61
 Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAATCCCAAGAGAGACAGAAATGCCAGAGCCAGCTCGAGGGCGCACTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCCTGCGAGCAATCTCTCAGAGATCCAGATCCAGATCCAGATTCATATGACGGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAGTGAACGAGTTTGAGAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
 QY 302 TGATGTGCGAGGATTGCAACAGATCATGAGAACACAGAGCCATAGTTGCGAGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
 Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACCACAGGCTTGCAGCTGGAGCTGCAAGTGGCGGCGGCGAGAGAC 466
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAsp 155
 RESULT 9
 ID ADO38314
 XX ADO38314 standard; peptide; 156 AA.
 AC ADO38314;
 XX 15-JUL-2004 (first entry)
 DT

XX Ara h 2 MHC-class II-presented epitope #1.
 DE Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
 XX Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
 KW Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;
 KW Antiallergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;
 KW MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
 KW influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
 KW rickettsia; rheumatoid arthritis; multiple sclerosis;
 KW lupus erythematosus; diabetes mellitus; myasthenia gravis;
 KW autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
 KW allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
 KW adenoma; peanut; Ara h 2.
 XX Arachis hypogaea.
 OS US2004058881-A1.
 PN 25-MAR-2004.
 XX 24-SEP-2002; 2002US-00253286.
 XX 24-SEP-2002; 2002US-00253286.
 PR (ANTI-) ANTIGEN EXPRESS INC.
 XX Humphreys RE, Xu M;
 XX WPI; 2004-294259/27.
 DR New non-naturally occurring protein or polypeptide modified by
 PT recombinant DNA techniques, useful for treating multiple sclerosis,
 PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
 PT colitis, cancer or psoriasis.
 XX Example 2; Page 18; 90pp; English.
 PS The invention relates to a non-naturally occurring protein or polypeptide
 XX (I) modified by recombinant DNA techniques comprising: a C-terminal
 CC element comprising an MHC Class II-presented epitope; an N-terminal
 CC element comprising an Ii-key motif; and an intervening element comprising
 CC a sequence of 4-11 amino acid residues where the modification by
 CC recombinant DNA techniques taking place within elements (b) and (c). Also
 CC described are methods for: suppressing or enhancing an immune response
 CC directed toward an MHC (major histocompatibility complex) Class II-
 CC presented epitope of interest. Suppressing an immune response directed
 CC toward an MHC Class II-presented epitope of interest comprises: providing
 CC a nucleic acid sequence encoding the MHC Class II-presented epitope of
 CC interest, the nucleic acid sequence encoding an Ii-key motif located 4-11
 CC amino acids upstream from the N-terminal residue of the MHC Class II-
 CC presented epitope of interest; and modifying the Ii-key motif to decrease
 CC its conformance to the archetypal Ii-Key regulatory motif. Enhancing an
 CC immune response directed toward an MHC Class II-presented epitope of
 CC interest comprises: providing a nucleic acid sequence encoding the MHC
 CC Class II-presented epitope of interest, the nucleic acid sequence lacking
 CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal
 CC residue of the MHC Class II-presented epitope of interest; and modifying
 CC the nucleic acid sequence to introduce an Ii-key motif appropriately
 CC spaced from the MHC Class II-presented epitope. The protein or
 CC polypeptide of interest corresponds to a protein or polypeptide encoded
 CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,
 CC preferably vaccinia virus. The non-naturally occurring protein or
 CC polypeptide (I) modified by recombinant DNA techniques is useful for
 CC treating infectious diseases caused or associated with infection by a
 CC bacterium, virus, parasite, fungus, rickettsia or other infectious
 CC agents. It is also useful for treating rheumatoid arthritis, multiple
 CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,
 CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,
 CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or
 CC adenomas. The present sequence represents the amino acid sequence of a
 CC peanut allergen Ara h 2 MHC class II-presented epitope used in the
 CC invention.

XX SQ Sequence 156 AA;

Alignment Scores:

Pred. No.:	4.23e-87	Length:	156
Score:	819.00	Matches:	153
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	64.74%	Indels:	0
DB:	8	Gaps:	0

US-10-728-323-2 (1-717) x ADO38314 (1-156)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCAG 61

DB 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23

QY 62 CAGTGGAACTCCAAGGAGACAGAAGATGCCAGGCAGCTCGAGGGCGAACCTGAGG 121

DB 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43

QY 122 CCTTGGCAGCAATCTCATGCGAGAGATCCAAAGTGCAGGATTCATATGAACGGAC 181

DB 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63

QY 182 CCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATCATCGAGGGCGCTGGA 241

DB 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83

QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301

DB 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103

QY 302 TGCATGTGCGAGCATTCGCAACAGATCATGGAGAACCGAGCCGATAGTTGCGAGGGAGG 361

DB 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123

QY 362 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGCGGCCCTTAGG 421

DB 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143

QY 422 GCACCAAGCGTTGCGACTTGGACGTGAAAGTGGCGGC 460

DB 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 10

ABU52576

ID ABU52576 standard; protein; 157 AA.

XX AC ABU52576;

XX DT 10-MAR-2003 (first entry)

XX DE Peanut Ara h2 mutant 5 (Q20A/Q31A/D60A/D67A) .

XX KW Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;

XX KM mutein; anaphylactic food allergen; antiallergenic; vaccine;

XX KW wound healing.

XX OS Homo sapiens.

XX OS WO200274250-A2.

XX PN 26-SEP-2002.

XX PD 18-MAR-2002; 2002WO-US009108.

XX PF 16-MAR-2001; 2001US-0276822P.

XX PR 18-MAR-2002; 2002US-00276822.

XX PA (PANA-) PANACEA PHARM.

XX XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;

PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

PI Rabjohn PA, Shin DS, Stanley JS;

XX WPI; 2003-018765/01.

XX New modified anaphylactic food allergen, useful for preventing or

PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX Example 13; Page; 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an

CC amino acid sequence that is substantially identical to that of natural

CC anaphylactic food allergen, except for a cysteine residue that has been

CC modified so that it cannot participate in the disulphide bond. The

CC modification may also comprise mutation of the IGE binding sites to

CC reduce allergenicity. Also included are: (1) a method of making a

CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding

CC or for causing a site specific mutation in the modified anaphylactic food

CC allergen; (3) a transgenic plant or animal expressing the modified

CC anaphylactic food allergen; (4) a method of treating an individual by

CC reducing the clinical response to a natural anaphylactic food allergen;

CC and an isolated fragment of peanut allergen Ara h 1. The modified

CC anaphylactic food allergen is useful for preventing or treating allergic

CC reactions associated with any natural allergen such as food, insect,

CC rubber or preferably anaphylactic allergens. It is also useful for

CC treating wounds in mammals such as bovine, canine, feline, caprine,

CC ovine, porcine, murine or equine species. The present sequence is a food

CC allergen, mutated to alter its IGE binding characteristics. Note: The

CC present sequence is not shown in the specification but was created by the

CC indexer using information provided in the specification

XX Sequence 157 AA;

SQ Alignment Scores:

Pred. No.:	1.65e-86	Length:	157
Score:	814.00	Matches:	153
Percent Similarity:	97.45%	Conservative:	0
Best Local Similarity:	97.45%	Mismatches:	4
Query Match:	64.35%	Indels:	0
DB:	6	Gaps:	0

US-10-728-323-2 (1-717) x ABU52576 (1-157)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCAG 61

DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgAla 20

QY 62 CAGTGGAACTCCAAGGAGACAGAAGATGCCAGGCAGCTCGAGGGCGCAACCTGAGG 121

DB 21 GlnTrpGluLeuGlnGlyAspArgCysAlaSerGlnLeuGluArgAlaAsnLeuArg 40

QY 122 CCTTGGCAGCAATCTCATGCGAGAGATCCAAAGTGCAGGATTCATATGAACGGAC 181

DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAla 60

QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGGCAGCGCTGGA 241

DB 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgGlyAlaGly 80

QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301

DB 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100

QY 302 TGCATGTGCGAGCATTCGCAACAGATCATGGAGAACCGAGCGATAGTTGCGAGGGAGG 361

DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120

QY 362 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGCGGCCCTTAGG 421

DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140

QY 422 GCACCAAGCGTTGCGACTTGGACGTGAAAGTGGCGGCAGACAGATAC 472

Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 11

ADMI2096
ID ADMI2096 standard; protein; 156 AA.

XX AC ADMI2096;

XX DT 20-MAY-2004 (first entry)

XX DE Arachis hypogaea 2 (Ara h2) protein.

XX KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-key;
KW MHC Class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.

XX OS Arachis hypogaea.

XX PN US2003235594-A1.

XX PD 25-DEC-2003.

XX PF 17-SEP-2002; 2002US-00245871.

XX PR 14-SEP-1999; 99US-00396813.

XX PR 17-JUL-2002; 2002US-00197000.

XX PA (ANTI-) ANTIGEN EXPRESS INC.

XX PI Humphreys R, Xu M;

XX DR WPI; 2004-070554/07.

XX PT Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.

XX PS Example 2; Page 19; 87pp; English.

XX CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian Ii-key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC represents a mammalian Ii key related protein of the invention.

XX SQ Sequence 156 AA;

Alignment Scores:
Pred. No.: 2,83e-86 Length: 156
Score: 812.00 Matches: 152
Percent Similarity: 99.35% Conservatives: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 64.19% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-2 (1-717) x ADMI2096 (1-156)

QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCACGCACTCTCGAGGACG	61
DB	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	23
QY	62	CAGTGGGAACCTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG	121
DB	24	GlnTrpGluLeuGlnGlnArgAspArgCysGlnSerGlnLeuGlnArgAlaAenLeuArg	43
QY	122	CCCTCGAGCAACATCTCATGCAAGATCCAAAGTCCAGTGAAGAGATTTCATATGAACGGAC	181
DB	44	ProCysGluGlyHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	63
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA	241
DB	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly	83
QY	242	TCCTCTCAGCACCAAGAGAGGGTGTTCGAATGAGCTGAACAGTTCGAGAACCAACCAAGG	301
DB	84	SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAenAsnGlnArg	103
QY	302	TGCATGTGCGAGGCATTGCAACAGATCATGAGTGGAGAACAGAGCGATAGTTGCGGGGAGG	361
DB	104	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	123
QY	362	CAACAGGACCAAGTTCAGAGGGAGCTCAGGAACCTTCCTCAACAGTCCGCGCTTAGG	421
DB	124	GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	143
QY	422	GCACCACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGC	460
DB	144	AlaProGlnArgCysAspLeuAspValGluSerGlyGly	156
RESULT 12			
ID	ABU52575	standard; protein; 157 AA.	
XX	AC	ABU52575;	
XX	DT	10-MAR-2003 (first entry)	
XX	DE	Peanut Ara h2 mutant 4 (E35A/P41A/D60A/D67A).	
XX	KW	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
XX	KW	muten; anaphylactic food allergen; anti-allergenic; vaccine;	
XX	KW	wound healing.	
XX	OS	Homo sapiens.	
XX	PN	WO200274250-A2.	
XX	PD	26-SEP-2002.	
XX	PF	18-MAR-2002; 2002WO-US009108.	
XX	PR	16-MAR-2001; 2001US-0276822P.	
XX	PR	18-MAR-2002; 2002US-00276822.	
XX	PA	(PANA-) PANACEA PHARM.	
XX	PI	Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;	
XX	PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;	
XX	PI	Rabjohn PA, Shin DS, Stanley JS;	
XX	DR	WPI; 2003-018765/01.	
XX	PT	New modified anaphylactic food allergen, useful for preventing or	
XX	PT	treating allergic reactions associated with e.g. anaphylactic allergens.	
XX	PS	Example 13; Page; 300pp; English.	
XX	CC	The invention relates to a modified anaphylactic food allergen has an	
XX	CC	amino acid sequence that is substantially identical to that of natural	
XX	CC	anaphylactic food allergen, except for a cysteine residue that has been	

CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IGE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IGE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 2,84e-86 Length: 157
Score: 812.00 Matches: 153
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 4
Query Match: 64.19% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABUS2575 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAGAGACAGAGATGCCAGAGCCAGCTCAGAGGGGGGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAACATCTCATGCAAGATCCAACTGACGAGGATTTCATATGAACGGGAC 181
Db 41 AlaCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAla 60
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCTCTCAGCACCAAGAGAGGTGTTCATGACTGAACTGAGTTTGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGCATTGCAACAGATCATGGAGAACCCAGAGCGATAGGTTGAGGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGCGTTGCCACTTGGACGTGCAAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 13
ABUS2577
ID ABUS2577 standard; protein; 157 AA.
XX
AC ABUS2577;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h2 mut. (W22A/E35A/P41A/D53A/D60A/D67A/R120A/L130A/L147A).
XX
KW Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;

KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX
OS Homo sapiens.
XX
PN WO200274250-A2.
XX
XX 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannion GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 13; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IGE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IGE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8,63e-82 Length: 157
Score: 774.00 Matches: 148
Percent Similarity: 94.27% Conservative: 0
Best Local Similarity: 94.27% Mismatches: 9
Query Match: 61.19% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABUS2577 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAGGAGACAGAGATGCCAGAGCCAGCTCAGAGGGGGGAACCTGAGG 121
Db 21 GlnAlaGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAACATCTCATGCAAGATCCAACTGACGAGGATTTCATATGAACGGGAC 181
Db 41 AlaCysGluGlnHisLeuMetGlnLysIleGlnArgAlaGluAspSerTyrGluArgAla 60
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241

Db 61 ProTyrSerProSerGlnAlaProTyr:SerProSerProTyrAspArgArgGlyAlaGly 80
Qy 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGGAGAACAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAenGluPheGluAenAsnGlnArg 100
Qy 302 TGCATGTGCGAGGCAATTGCCACAGATCATGAGNACCAAGAGCGATAGTTGCCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArgLeuGlnGlyAla 120
Qy 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
Db 121 GlnGlnGlnGlnGlnPheLysArgGluAlaAargAsnLeuProGlnGlnCysGlyLeuArg 140
Qy 422 GCACCAAGCGTTCCGACTTGGACGTGCAAGTGGCGGCGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 14
ID AAY40968 standard; protein; 166 AA.
AC AAY40968;
XX
XX
DT 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX
XX
XX Recombinant Ara h 2 protein sequence.
XX
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
XX allergic reaction; Ara h 2.
XX
XX Arachis hypogaea.
XX
XX WO9945961-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US005494.
XX
XX 12-MAR-1998; 98US-0077763P.
PR 11-MAR-1999; 99US-00077763.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX
XX WPI; 1999-551218/46.
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
XX
XX Disclosure; Page 81; 193pp; English.
XX
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents a recombinant Ara h 2 protein. (Updated on 17
CC -OCT-2003 to standardise OS field)
XX
XX Sequence 166 AA;
SQ

Alignment Scores:
Pred. No.: 2e-81 Length: 166
Score: 771.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-2 (1-717) x AAY40968 (1-166)
Qy 50 TCCTCCAGGCGCAGCAGTGGGAACCTCCAGGAGAGACAGAGATGCCAGAGCCAGCTCGAGAG 109
Db 15 SerAlaAargGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArg 34
Qy 110 GCCAACTCAGGCGCTCCGAGCAACATCTCATCAGAGAAGATCCAACTGCAGGAGATTCA 169
Db 35 AlaAsnLeuAargProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer 54
Qy 170 TATGAACGGGACCCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG 229
Db 55 TyrGluAArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 74
Qy 230 AGAGGCGCTGCATCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAenGluPheGlu 94
Qy 290 AACAAACCAAGGTGCATGTGCCAGGCAATTCGAACAGATCATGAGAACCAAGAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArg 114
Qy 350 TTCAGGCGGAGGCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAG 409
Db 115 LeuGlnGlyArgGlnGlnGlnPheLysArgGluLeuAenGluLeuProGlnGln 134
Qy 410 TGGCGCTTAGGCGCACAGCGTTGGAGCTTGGAGCTGCAAGTGGCGGCGAGAGACAGA 469
Db 135 CysGlyLeuAArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 154
Qy 470 TAC 472
Db 155 Tyr 155
RESULT 15
ABU52482
ID ABU52482 standard; protein; 166 AA.
XX
XX AC ABU52482;
XX
XX 10-MAR-2003 (first entry)
XX
XX Peanut Ara h2 with N-terminal a T7 tag and a C-terminal His tag.
DE
XX
XX Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;
KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
XX Arachis hypogaea.
OS Synthetic.
OS Enterobacteria phage T7.
XX
XX WO200274250-A2.
XX
XX 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US0009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NS, Kopper RA, Maleki SJ;
XX Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT

PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX
PS
XX Example 13; Fig 51; 300pp; English.

XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen, Ara h2, with an N-terminal 17 tag and a C-terminal His
CC tag

XX
SQ Sequence 166 AA;

Alignment Scores:
Pred. No.: 2e-81 Length: 166
Score: 771.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-171) x ABUS2482 (1-166)

Qy	50	TCGCGAGGCGAGCTGGGAATCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGG	109
Db	15	SerAlaArgGlnGlnTrpGluLeuGlnGlnArgCysGlnSerGlnLeuGluArg	34
Qy	110	GCGAAGCTGAGGCGCTCGGAGCAACATCTCATCAGAGATCCAAAGTGCAGGAGTTCA	169
Db	35	AlaAsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer	54
Qy	170	TATGAACGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG	229
Db	55	TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg	74
Qy	230	AGAGGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAG	289
Db	75	ArgGlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu	94
Qy	290	AACACCAAGGTGCATGTGCGAGGCATTTCACAGATCATGAGAACCGAGCGGATAGG	349
Db	95	AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg	114
Qy	350	TTGCGGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAG	409
Db	115	LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln	134
Qy	410	TGCGGCTTAGGGCACACAGCGTTGCGACTTGAGCTCGAAAGTGGCGGCAGAGACAGA	469
Db	135	CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg	154
Qy	470	TAC 472	
Db	155	Tyr 155	

Search completed: August 24, 2005, 09:55:57
Job time : 43.6071 sec8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 07:15:51 ; Search time 8.97718 Seconds
(without alignments)
15369.504 Million cell updates/sec

Title: us-10-728-323-2

Perfect score: 1265

Sequence: 1 gctcacatagtagcc.....taaaagatcatgtttgtt 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US1078323/runat_23082005_124355_29222/app_query.fasta_1.4757
-DB=PIR_79 -QMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40 cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORW=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1078323@cgn_1_1_76@runat_23082005_124355_29222 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	278	22.0	153	2 A33090	conglutin delta pr
2	243	19.2	158	1 T05710	2S albumin precurs
3	187.5	14.8	80	2 B23617	conglutin delta-2
4	184.5	14.6	323	2 S38887	2S albumin - commo
5	173.5	13.7	141	2 T10257	2S albumin precurs
6	154	12.2	154	2 S14947	2S albumin - Brazi
7	151	11.9	295	2 S01062	2S seed storage pr
8	147	11.6	186	2 A29802	napin precursor (g
9	146.5	11.6	164	1 NWMU1	2S albumin I precu
10	140.5	11.1	146	2 S14946	2S seed storage pr
11	139	11.0	155	2 JCS379	mablin II precu
12	139	11.0	258	1 RZCS	2S seed storage pr
13	132.5	10.5	165	2 T09252	seed storage prote
14	129	10.2	170	2 T08012	2S seed storage pr

15	128	10.1	172	2 S18871	2S-like storage pr
16	126.5	10.0	70	2 A59346	seed storage prote
17	125	9.9	162	2 T08013	2S seed storage pr
18	124.5	9.8	162	2 S49259	albumin 4 - easter
19	122.5	9.7	162	2 T08010	2S seed storage pr
20	121	9.6	100	2 S48180	mablin IV - Yun
21	121	9.6	174	2 PS0425	napin B3 precursor
22	119.5	9.4	180	2 S52025	napin (clones BMMN
23	118.5	9.4	106	2 S26636	napin n1b - rape
24	118.5	9.4	110	2 S20350	napin n1a - rape
25	118	9.3	104	2 S48176	mablin I-1 - Yun
26	117	9.2	104	2 S48178	mablin III - Yun
27	115	9.1	152	2 PS0427	napin AH1 precursor
28	111.5	8.8	173	2 T08011	2S seed storage pr
29	108.5	8.6	178	2 S07828	napin B - rape
30	108.5	8.6	178	2 S25127	2S storage protein
31	107	8.5	139	2 T09850	albumin 2S storage
32	107	8.5	139	2 T09878	albumin 2S storage
33	106.5	8.4	178	1 NMRP2	napin 2 precursor
34	106.5	8.4	178	2 A25997	napin precursor (n
35	105.5	8.3	164	1 NWMU3	2S albumin 3 precu
36	105.5	8.3	180	2 S10018	napin (clone BngNA
37	105.5	8.3	1099	2 A56155	tumor suppressor p
38	105.5	8.3	1909	2 A45592	liver stage anticp
39	104.5	8.3	1363	2 T15496	hypothetical prote
40	104	8.2	145	2 S65479	allergen Sin a I (
41	103.5	8.2	166	1 NWMU4	2S albumin 4 precu
42	103.5	8.2	178	2 S25130	2S storage protein
43	102.5	8.1	178	2 S25134	2S storage protein
44	101.5	8.0	178	2 S25137	2S storage protein
45	101	8.0	145	2 PC1246	Sin a I allergen 1

ALIGNMENTS

RESULT 1

A33090
conglutin delta precursor - narrow-leaved blue lupine
C:Species: Lupinus angustifolius (narrow-leaved blue lupine)
C>Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 09-Jul-2004
C:Accession: S12404; A33090
R:Gayler, K.R.; Kolivas, S.; MacFarlane, A.J.; Lilley, G.G.; Baldi, M.; Blagrove, R.J.;
Plant Mol. Biol. 15, 879-893, 1990
A:title: Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta,
A:Reference number: S12404; MWID:91355912; PMID:2103479
A:Accession: S12404
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-153 <GAY>
A:Cross-references: UNIPROT:Q99235; GB:X53523; NID:g19140; PIDN:CAA37598.1; PID:g19141
C:Superfamily: soybean 2S albumin
C:Keywords: seed
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-59/Product: conglutin small chain #status predicted <MAT1>
F:73-153/Product: conglutin large chain #status predicted <MAT2>

Alignment Scores:

Pred. No.:	3,84e-18	Length:	153
Score:	278.00	Matches:	62
Percent Similarity:	55.00%	Conservative:	26
Best Local Similarity:	38.75%	Mismatches:	46
Query Match:	21.98%	Indels:	26
DB:	2	Gaps:	5

US-10-728-323-2 (1-717) x A33090 (1-153)

Qy	2	CTCACCATAGTAGCCCTC---GCCCTTTTCCTCTCGTCCCGCCACGCATCTCGAGG 58
Db	4	LeuThrIleuLeuValAlaLaLeuValLeuValValHisThrSerAlaPhe 23
Qy	59	CAGCAGTGGGAATCCCAAGGAGACAGATGCAGAGCCAGCTCCGAGAGGGCGAACCTTG 118
Db	24	Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnValAsnLeu 39

```
QY 119 AGCCCTGCAGCAACATCTCATGCGAAGATCCAAGTCGACGAGGATTCATATGAACGG 178
Db 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGlnGluGluAsp 59
QY 179 GACCCGTACAGC-----CCTAGTCAGGATCCG 205
Db 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerGlnGlu 79
QY 206 TAGAGCCCTAGTCATATGATCGGAGGGCGCTGATCTCTCAGCACCAAGAGAGGTGT 265
Db 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCAATGAGTGAACGAGTTTGAGAACAAAGGTGCGATGTCGAGGCGATTGCCAACAG 325
Db 91 CysGluGlnLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGGAGAACGAGCGGATAGTTGCGAGGAGGCAACAGAGCAACAGTTCAAGAGG 385
Db 110 IleTyrGluSerGlnSerGluGlnCysGluGlySerGlnGlnGlnGlnGluGln 129
QY 386 GAGCTCAGGAACTTGCTCAACAGTGGCGCTTAGGGCACCAGCGTTGCGACTTGGAC 445
Db 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgCysAspValAsn 149

RESULT 2
T05710
2S albumin precursor - soybean
N;Alternate names: aspartic acid-rich peptide
C;Species: Glycine max (soybean)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: T05710; A28485
R;Wang, J.; Pichersky, E.
Plant Physiol. 114, 1567, 1997
A;Title: Nucleotide Sequence of S-Adenosyl-L-Methionine:Caffeic Acid 3-O-Methyltransferase
A;Reference number: Z15424
A;Accession: T05710
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-158 <W>N>
A;Cross-references: UNIPROT:P19594; EMBL:AF005030; NID:g2305019; PID:g2305020
A;Experimental source: cultivar Hodgson 78; cotyledon
R;Odani, S.; Koide, T.; Ono, T.
J. Biol. Chem. 262, 10502-10505, 1987
A;Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(L
A;Reference number: A28485; MUID:87280104; PMID:3611081
A;Accession: A28485
A;Molecule type: protein
A;Residues: 22-64 <ODA>
C;Superfamily: soybean 2S albumin
C;Keywords: glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Product: 2S albumin small chain #status predicted <WAT>
F;54-56/Region: cell attachment (R-G-D) motif
F;39/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 7,86e-15 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x T05710 (1-158)

QY 5 ACCATAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGGAGCGACG 64
Db 5 ThrIleLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGGACACAGAAATGCCAGAGCGAGCTCGAGAGGGCGGAACCTGAGGCC 124
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
```

```
QY 125 TGCAGCAACATCTCATGCGAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GATCCTCTCTCAG-----CACCAGAGAGCGTGTGCAATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGGAGAACAAACAAAGGTGCGATGTCGAGGCGATTGCAACAGATCATGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 ACGATAGATTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGGCGCTTAGGGCACCACAGCGTTGCGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
B23617
conglutin delta-2 large chain - narrow-leaved blue lupine
C;Species: Lupinus angustifolius (narrow-leaved blue lupine)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: B23617
R;Lilley, G.G.; Inglis, A.S.
FEBS Lett. 195, 235-241, 1986
A;Title: Amino acid sequence of conglutin delta, a sulfur-rich seed protein of Lupinus ar
A;Residues: 1-80 <LIL>
A;Cross-references: UNIPROT:P09931
C;Superfamily: soybean 2S albumin

Alignment Scores:
Pred. No.: 1.42e-09 Length: 80
Score: 187.50 Matches: 34
Percent Similarity: 73.13% Conservative: 15
Best Local Similarity: 50.75% Mismatches: 17
Query Match: 14.82% Indels: 1
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x B23617 (1-80)

QY 245 TCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGC 304
Db 11 SerGluGluLeuAspGlnCysGluGlnLeuAsnGluLeu---AsnSerGlnArgCys 29
QY 305 ATGTGCGAGCGCATTCGAACAGATCATGAGACCAAGCGATAGGTTCGAGGGGAGGCAA 364
Db 30 GlnCysArgAlaLeuGlnGlnIleTyrGluSerGlnSerGluGlnCysGluGlyArgGln 49
QY 365 CAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCCTTAGGGCA 424
Db 50 GlnGluGlnGlnLeuGluGlyGluLeuGluLysLeuProArgIleCysGlyPheGlyPro 69
QY 425 CCACAGCGTTGCGACTTGGAC 445
Db 70 LeuArgArgCysAsnIleAsn 76

RESULT 4
S38887
2S albumin - common sunflower (fragment)
C;Species: Helianthus annuus (common sunflower)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S38887
```

R;Thoyts, P.J.E.; Millichip, M.; Stobart, A.K.; Griffiths, W.T.; Napier, J.A.; Shewry, P.
submitted to the EMBL Data Library, November 1993

A;Reference number: S3887
A;Accession: S3887
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-323 <THO>
A;Cross-references: UNIPROT:Q3928; EMBL:X76101; NID:g429181; PID:g429182
C;Superfamily: gliadin

Alignment Scores:
Pred. No.: 2.65e-09 Length: 323
Score: 184.50 Matches: 46
Percent Similarity: 48.99% Conservatives: 27
Best Local Similarity: 30.87% Mismatches: 43
Query Match: 14.58% Indels: 33
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x S3887 (1-323)

```
QY 83 AGAAGATGC---CAGAGCCAGCTCGAGAGCGCAACCTGAGGCCCTCGGAGCAACATCTC 139
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 GlnGlnCysArgGlnSerGluLeuGlnArgPro---ValSerGlnCysGlnArgTy-Val 195
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 ATGCAGAGATCCAA-----CGTGACGAG 163
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 GluGlnGlnIleGlnSerSerArgProTyrglnGlnSerProTyrglnGlnGln 215
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 GATTTCATATGAACGGGACCCGCTACAGCCCT-----AGTCAG 199
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 216 SerProTyrglnSerArgGlnGlnSerProTyrglnGlnArgGlnGlnProTyrglnGln 235
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 GATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAA--- 256
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 ArgProTyrglnGlnArgProTyrglnGlnArgGlyArgGlnGlnGlnGlnGlnGly 255
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 ---CAGAGGTGTTCAATGAGCTGAACAGAGTTTCAGAACCAACAAAGGTGATGTGCGAG 313
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 LeuGlnGlnCysCysAsnGluLeuGlnAsnVal-----ArgArgGluCysGlnCysGlu 273
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 GCATTGCAACAGATC-----ATGGAGAACAGAGCGATAGTTGTCAG 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 274 AlaIleGlyGluValGlyGlnArgMetArgGlnGlnGlnGlnGlnGlnArgGlnTy 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 GGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTCACACAGTGCAGC 415
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 GlyGlyGlnGlnThrGlnThrValGluArgIleLeuGluAsnLeuProAsnGlnCysAsp 313
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 CTTAGGCGACCAACAGCGCTTGGACTTG 442
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 314 LeuAspVal---GlnGlnCysAsnIle 321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 5

T10257
2S albumin precursor - cucurbit
N;Alternate names: prepro2S albumin
C;Species: Cucurbita sp. (cucurbit)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10257; S19323
R;Hara-Nishimura, I.; Takeuchi, Y.; Inoue, K.; Nishimura, M.
Plant J. 4, 793-800, 1993
A;Title: Vesicle transport and processing of the precursor to 2S albumin in pumpkin.
A;Reference number: Z17000; MUID:94100993; PMID:8275099
A;Accession: T10257
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA; protein
A;Residues: 1-141 <HAR1>
A;Cross-references: EMBL:D16560; NID:g459404; PIDN:BA003993.1; PID:g459405
A;Experimental source: seed, storage deposition stage; cotyledon
A;Note: soluble seed protein
R;Hara-Nishimura, I.; Inoue, K.; Nishimura, M.
FEBS Lett. 294, 89-93, 1991
A;Title: A unique vacuolar processing enzyme responsible for conversion of several propu

A;Reference number: S19323; MUID:92077151; PMID:1743299

A;Accession: S19323
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 26-45;65-84 <HAR2>
C;Keywords: seed; storage protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-35/Domain: propeptide #status experimental <PRO>
F;36-141/Product: 2S albumin #status experimental <MAT>

Alignment Scores:
Pred. No.: 2.97e-08 Length: 141
Score: 173.50 Matches: 52
Percent Similarity: 49.37% Conservatives: 26
Best Local Similarity: 32.91% Mismatches: 47
Query Match: 13.72% Indels: 33
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x T10257 (1-141)

```
QY 2 CTCACCATACTAGTAGCCCTC---GCCCTTTTCTCTCTCGCTGCCACGCACTCTGCGAGG 58
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 LeuThrSerIleIleAlaLeuPheAlaValAlaLeuValAlaAspAlaTyraIaTy 23
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 CAG-----CAGTGGGAACCTCCAAGCA---GACAGAAGATGC 91
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 ArgThrThrIleThrThrValGluValGluValAsnArgGlnGlyArgGluGluArgCys 43
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 CAGAGCCAGCTCGAGAGGCGCAACCTGAGGCCCTCGAGCAACATCTCATGCAGAGATC 151
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 44 ArgGlnMetSerAlaArgGluGluLeuArgSerCysGluGlnTyrluArgGlnGln--- 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 CAACGTGACGAGGATTCATATGAACGGGACCCCTACAGCCCTAGT-----CAGGATCCG 205
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 -----SerArgAspValLeuGlnMetArgGlyIleGluAsnPro 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 Trp-----ArgArgGluGlyGlySerPhe-----AspGluCys 86
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 TGCATGAGCTGAACAGTGTGAGAACCAACCAAGGTGCATGTGCGAGGCATTCGACACAG 325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 CysArgGluLeuLysAsnValAsp-----GluGluCysArgCysAspMetLeuGluGlu 104
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 ATCATCGAGAACCAACGAGCCGATAGTTGCGAGGGAGGCAACAGGACCAACAGTTTCAAGAGG 385
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 IleAlaArgGluGlnArgGlnAlaArgGly---GlnGluGlyArgGlnMetLeuGln 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 GAGCTCAGGAACCTGCCTCAACAGTGGCGCTTAGGCGCACCAAGCGTTGCGAC 439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 LysAlaArgAsnLeuProSerMetCysGlyIleArg---ProGlnArgCysAsp 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 6

S14947
2S albumin - Brazil nut
C;Species: Bertholletia excelsa (Brazil nut)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C;Accession: S14947
R;Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Grossi
Plant Mol. Biol. 16, 437-448, 1991
A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin fr
A;Reference number: S14946; MUID:91370890; PMID:1840683
A;Accession: S14947
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <CAN>
A;Cross-references: EMBL:X54491; NID:g17712; PIDN:CAA38363.1; PID:g17713
C;Genetics:
A;Introns: 67/3
C;Superfamily: wheat alpha-amylase inhibitor
Alignment Scores:
Pred. No.: 2.07e-06 Length: 154

Score:	154.00	Matches:	45
Percent Similarity:	25.34%	Conservative:	28
Best Local Similarity:	27.95%	Mismatches:	52
Query Match:	12.17%	Indels:	36
DB:	2	Gaps:	10

US-10-728-323-2 (1-717) x S14947 (1-154)

Qy	8	ATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCAGCATCTGGC-----	55
Db	7	ValAlaAlaLeuLeuAlaLeuLeuValLeuGlnAlaThrAlaPheArgThr	26
Qy	56	-----AGGCAGCAGTGGGAACCTCCAAG-----GACACAGA	88
Db	27	ValThrThrLeuGluGluGluGluAsnProArgGlyArgSerGluGlnGln	46
Qy	89	TGCACAGGCAGCTCAGAGG---CGGAACCTGAGGCCCTCGGAGCAACATCTCATCGAG	145
Db	47	CysArgGluGlnMetGluArgGlnGlnGlnLeuAsnHisCysArgMetTyrLeuArgGln	66
Qy	146	AAGATCCAACGCTGACGAGGATTCATATAACGGGACCGTCAGCCCTAGTCAGGATCCG	205
Db	67	GlnMet-----GluGluSerProTyrGln-----	74
Qy	206	TACAGCCCTAGTCCATATGATCGAGAGGCGCTGTGATCTCTCAGACCAAGAGAGGTGT	265
Db	75	---AsnProArgProLeuArgArg-----GlyGluGluProHisLeuAspGluCys	90
Qy	266	TGCAATGAGCTGAACGAGTTTCAGAACCAACAAAGGTGCATGTCGGAGCATTCACAAG	325
Db	91	CysGluGlnLeuGluArgMetAsp-----GluMetCysArgCysGluGlyLeuArgMet	108
Qy	326	ATCATGGAGAACCCAGAGCGAT-----AGTTGTCAGGGGAGCGCAACAGGACCAACAGTTC	379
Db	109	MetLeuArgArgGlnArgGluGluMetGluLeuGlnGly---GluGlnMetGlnArgIle	127
Qy	380	AACAGGGAGCTCAGGAACCTTGCTCAACAGTGCGCCCTTAGGGCACCAACAGCGTTGGAC	439
Db	128	MetArgGlyAlaGluAsnLeuLeuSerArgCysAsnLeu---SerProGlnArgCysPro	148
Qy	440	TTG	442
Db	147	Met	147

```

RESULT 7
S01062
2S seed storage protein precursor (clone HaGs) - common sunflower
N;Alternate names: 2S albumin storage protein
C;Species: Helianthus annuus (common sunflower)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: S01062
R;Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Thod
Mol. Gen. Genet. 210, 211-218, 1987
A;Title: Sequence and expression of a gene encoding an albumin storage protein in sunflo
A;Reference number: S01062; MUID:88142538; PMID:2830455
A;Accession: S01062
A;Molecule type: DNA
A;Residues: 1-295 <ALL>
A;Cross-references: UNIPROT:P15461; EMBL:X06410; NID:gl18799; PIDN:CAA29699, 1; PID:gl18800
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:
A;Introns: 192/2
C;Superfamily: gliadin
F;1-20/Domain: signal sequence #status predicted <SIG>
F;162-295/Product: 2S seed storage protein #status experimental <MAT>

```

Alignment Scores:	
Pred. No.:	3,93e-06
Score:	151.00
Percent Similarity:	40.39%
Best Local Similarity:	23.15%
Query Match:	11.94%
DB:	2
Gaps:	8
Indels:	56
Mismatches:	65
Conservative:	35
Matches:	47
Length:	295

US-10-728-323-2 (1-717) x S01062 (1-295)

QY 8 ATACTAGTAGCCCTCGCCCTTTTCCTCCTGCTGCCACGCATCTGG-----55
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 5 IleValLeuAlaAlaPheAlaIleValAlaPheAlaThrAlaHisThrIle24 24
QY 56 -----AGGCAGCATGGGAATCCAAGGACAGACAAGAATGCCAGCGCAG 100
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 25 IieThrThrThrIleGluAspGluAsnProIleSerGlyGlnArgInValSerGlnArg 44
QY 101 CTCGAGAGGCGCAACTCAGGCCCTCGCGCAACATCTCATG-----CAGAAGATCCAA 150
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 45 IieGlnGlyGlnArgLeuAsnGlnCysArgMetPheLeuGlnGlnGlyGlnAsnIlePro 64
QY 155 CGTGACGAGGATTTCATATCAACGGGACCCTACAGCCCTAGTTCAGGATCCGTACAGCCCT 214
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 65 ArgGluPheAspAsn-----69
QY 215 AGTCCATATGATCGAGAGGCGCTGGATCCTCTCAGACCNAGAGAGAGTGTTCATGAG 274
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 70 -----ProGlnMetGlyArgGlnGlnGlnGlnLeuGlnGlnCysCysGlnGlu 86
QY 275 CTGAACAGCTTTTGACAACAACCAAAGGTGCATGTGCAGGCATTGCAACAGATCATCGGAG 334
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 87 LeuGlnAsnIleGluGly-----GlnCysGlnCysGluAlaValLysGlnValPheArg 104
QY 335 AACACAGAGCATAGTTGTCAG-----GGGAGG 361
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 105 GluAlaGlnGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 124
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 125 GlnGlnThrGlnGlnLeuLysGlnLysAlaGlnIleLeuProAsnValCysAsnLeuGln 144
QY 422 GCACACACAGCTGTGCAGTTG-----GACGTCGAA 451
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 145 Ser--ArgArgCysGluIleGlyThrIleThrThrThrValThrGluSerAsnIleAsp 163
QY 452 AGTGGCGGCAGACAGATACTAAACACTTATCAAAAAAAGAAAAGAAAAGAA 511
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 164 IleProPheArgAspArg-----ProPheGlyThrGlySerGlnGlnCysArgGlu 180
QY 512 AATAGCTTA 520
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 181 ThrGluIle 183

RESULT 8
A29802
napin precursor (gNa) - rape
N;Alternate names: 1.7 S seed storage protein
C;Species: Brassica napus (rape)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 09-Jul-2004
C;Accession: A29802
R;Scotfield, S.R.; Crouch, M.L.
J. Biol. Chem. 262, 12202-12208, 1987
A;Title: Nucleotide sequence of a member of the napin storage protein family f
A;Reference number: A29802; PMID:87308225; PMID:3040733
A;Accession: A29802
A;Molecule type: DNA
A;Residues: 1-186 <SCO>
A;Cross-references: UNIPROT:P09893; GB:J02782; NID:g1617178; PIDN:AAA33007.1; P
C;Genetics:
C;Superfamily: wheat alpha-amylase inhibitor

Alignment Scores:

Pred. No.:	Score:	Length:
A29802	147.00	186
Percent Similarity:	40.98%	Matches: 47
Best Local Similarity:	25.68%	Conservative: 28
Query Match:	1-162*	Mismatches: 52
DB:	2	Indels: 56
		Gaps: 9


```

Db      105 TyrAspSerPheMetAspSerAspSerGlnGluAspAlaProLeuAsnGlnArgArgArg 124
Qy      350 TTG-----CAGGGGAGGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTG 400
Db      125 ArgArgGluGlyArgGlyArgGluGluGluAlaMetGluArg---AlaAlaTyrLeu 143
Qy      401 CCTCAACAGTCGGCCTTAGG---GCACCACAGCGTTGGGACTTGGACGTCGAAAGT 454
Db      144 ProAsnThrCysAsnValArgGluProProArgArgCysAspIleGlnArgHisSer 162

```

Search completed: August 24, 2005, 10:06:37
Job time : 11.9772 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 43.8791 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctccacatactagtagcccc.....taaaagatcatgttttgg 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10728323/runat_23082005_124355_29210/app_query.fasta.1.4757
-DB=Uniprot 03 -QFMT=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR=score=pept -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1 291 @runat_23082005_124355_29210 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	842	66.6	179	2	Q647H0 arachis hyp
2	838	66.2	175	2	Q6PSU1 arachis hyp
3	819	64.7	156	2	Q941R0 arachis hyp
4	815	64.4	172	2	Q8GV20 arachis hyp
5	815	64.4	187	2	Q6PSU2 arachis hyp
6	808	63.9	169	2	Q7Y1C0 arachis hyp
7	792	62.6	166	2	Q84TU1 arachis hyp
8	454.5	35.9	145	2	Q647G9 arachis hyp
9	443.5	35.1	144	2	Q8W251 arachis hyp
10	388.5	30.7	129	2	Q9SGQ5 arachis hyp
11	300	23.7	160	2	Q9SOH1 arachis hyp
12	299	23.6	158	2	Q647G8 arachis hyp
13	278	22.0	153	2	Q99235 lupinus ang
14	243	19.2	158	1	P19594 glycine max
15	241.5	19.1	155	2	Q92NZ4 glycine max
16	187.5	14.6	80	1	CG2L_LUPAN lupinus ang

17	184.5	14.6	323	2	Q39928 helianthus
18	177.5	14.0	168	2	Q8LPD3 linum usita
19	173.5	13.7	141	2	Q39649 cucurbita c
20	167	13.2	160	2	Q71HN1 ficus awkeo
21	164.5	13.0	148	2	Q9XHP1 sesamum ind
22	159.5	12.6	143	2	Q8VX62 sesamum ind
23	159	12.6	161	2	Q7Y1C2 juglans nig
24	153.5	12.1	139	2	P93198 juglans reg
25	153	12.1	153	2	Q9AUD1 sesamum ind
26	151	11.9	295	1	2SS5 HELAN
27	149.5	11.7	167	2	Q84JW2 vitis vinif
28	148.5	11.7	140	2	Q8L694 momordica c
29	148	11.7	285	2	Q8GUD8 helianthus
30	147	11.6	186	1	2SS5 BRANA
31	146.5	11.6	164	1	2SS1 ARATH
32	144.5	11.4	165	2	Q9FHP31 arabidopsis
33	144	11.4	169	2	Q8LPD4 linum usita
34	143.5	11.3	179	2	Q84NG9 vitis vinif
35	141.5	11.2	138	2	Q8H2B8 anacardium
36	141	11.1	126	2	Q8RW54 helianthus
37	140.5	11.1	146	1	2SS BEREX
38	139	11.0	155	1	2SS2 CAPMA
39	139	11.0	184	2	Q42444 brassica ca
40	139	11.0	258	1	2SS RICCO
41	132.5	10.5	165	2	Q40850 picea glauc
42	132.5	10.5	167	2	Q81411 picea glauc
43	130.5	10.3	161	2	Q81410 picea glauc
44	129.5	10.2	173	2	Q81412 picea glauc
45	129	10.2	170	2	Q64931 pseudotsuga

ALIGNMENTS

RESULT 1

Q647H0 PRELIMINARY; PRT; 179 AA.
AC Q647H0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE 2S protein 1 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY22689; AAU21494.1; -;
FT NON_TER 1
SQ SEQUENCE 179 AA; 20850 MW; 68CE6453B409E1DC CRC64;

Alignment Scores:

Pred. No.: 2,57e-65 Length: 179
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q647H0 (1-179)

QY	2	CTCACCATAGTAGTACCTCGCCCTTTTCTCTCGCTGCCACGCATCTGGAGGCGAG 61
Db	23	LeuThrIleLeuValAlaLeuAlaLeuLeuLeuAlaHisAlaSerAlaArgIn 42
QY	62	CAGTGGGAAGTCCCAAGGAGACAGAGTGCAGAGCGAGCTCGAGAGGGGGAACCTGAGG 121
Db	43	GlntrpGluLeuGlnGlyAspArgGcysGlnSerGlnLeuGluArgAlaAsnLeuArg 62

QY	122	CCCTGCGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	63	ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp	82
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATGATCGAGAGCGCTGGA	241
Db	83	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	102
QY	242	TCCTCTCAGACCAAGAGAGGTGTCATGACCTGAACGAGTTTGAACAACCAAGG	301
Db	103	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	122
QY	302	TGCATGTCCGAGGCAATTCGAACAGATCATCGAGAACCCAGAGCGATAGTTTCAGGGGAGG	361
Db	123	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	142
QY	362	CAACAGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTCGCTCAACAGTGCAGCGCTTAGG	421
Db	143	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	162
QY	422	GCACACAGCGTTCCGACTTGGAGCTTGGAGTCCGAAGTGGCGGCAGACAGATAC	472
Db	163	AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr	179
RESULT 2			
ID	Q6PSU1	PRELIMINARY;	PRT; 175 AA.
AC	Q6PSU1; 2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Seed storage protein SSP2 (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]_TaxID=3818;		
RP	SEQUENCE FROM N.A.		
RA	Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY581854; AAT00599.1; -.		
DR	InterPro; IPR003612; AAI.		
DR	Pfam; PF00234; Tryp_alpha_amyl; 1.		
DR	SMART; SM00499; AAI; 1.		
FT	NON_TER 1 1		
SQ	SEQUENCE 175 AA; 20463 MW; B3FB5806B70733DE CRC64;		
Alignment Scores:			
Pred. No.:	5.77e-65	Length:	175
Score:	838.00	Matches:	156
Percent Similarity:	99.36%	Conservative:	0
Best Local Similarity:	99.36%	Mismatches:	1
Query Match:	66.25%	Indels:	0
DB:	2	Gaps:	0
US-10-728-323-2 (1-717) x Q6PSU1 (1-175)			
QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTCCGCCACGCATCTCGAGGCAG	61
Db	19	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	38
QY	62	CAGTGGGAACCTCAAGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGGAACTGAGG	121
Db	39	GlnTrpGluPheGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	58
QY	122	CCCTGCGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	59	ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp	78
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241

Db	79	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	98
QY	242	TCCTCTCAGACCAAGAGAGGTGTCATGACCTGAACGAGTTTGAACAACCAAGG	301
Db	99	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	118
QY	302	TGCATGTCCGAGGCAATTCGAACAGATCATGGAGAACCCAGAGCGATAGTTTCAGGGGAGG	361
Db	119	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	138
QY	362	CAACAGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTCGCTCAACAGTGCAGCGCTTAGG	421
Db	139	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	158
QY	422	GCACACAGCGTTCCGACTTGGAGCTTGGAGTCCGAAGTGGCGGCAGACAGATAC	472
Db	159	AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr	175
RESULT 3			
ID	Q941R0	PRELIMINARY;	PRT; 156 AA.
AC	Q941R0; 2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Allergen II (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]_TaxID=3818;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21192158; PubMed=11295663; DOI=10.1067/mai.2001.113522;		
RA	Viquez O.M., Summer C.G., Dodo H.W.;		
RT	"Isolation and molecular characterization of the first genomic clone		
RT	of a major peanut allergen, Ara h 2.;"		
RL	J. Allergy Clin. Immunol. 107:713-717(2001).		
DR	EMBL; AY007229; AAK96887.1; -.		
DR	InterPro; IPR003612; AAI.		
DR	Pfam; PF00234; Tryp_alpha_amyl; 1.		
DR	SMART; SM00499; AAI; 1.		
FT	NON_TER 156 156		
SQ	SEQUENCE 156 AA; 18167 MW; 664D6832F91F36D1 CRC64;		
Alignment Scores:			
Pred. No.:	2.68e-63	Length:	156
Score:	819.00	Matches:	153
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	64.74%	Indels:	0
DB:	2	Gaps:	0
US-10-728-323-2 (1-717) x Q941R0 (1-156)			
QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTCCGCCACGCATCTCGAGGCAG	61
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	23
QY	62	CAGTGGGAACCTCAAGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGGAACTGAGG	121
Db	24	GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	43
QY	122	CCCTGCGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	44	ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp	63
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	83

QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
 Db 84 SerSerGlnHisGlnGluAraGysCysAsnGluLeuAsnGluPheGluAsnGlnAraG 103
 QY 302 TGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGCGATAGTTTGCAGGGGAGG 361
 Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyAraG 123
 QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCTCAACAGTCGGCGCTTAGG 421
 Db 124 GlnGlnGlnGlnPheGlyAraGLeuAraGAsnLeuProGlnGlnCysGlyLeuAraG 143
 QY 422 GCACCACAGCGTTGCGACTTGGAGCTGCGAAAGTGGCGGC 460
 Db 144 AlaProGlnAraGysAspLeuAspValGluSerGlyGly 156
 RESULT 4
 Q8GV20 PRELIMINARY; PRT; 172 AA.
 AC Q8GV20; 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Allergen Ara h 2.02.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chatel J.-M., Bernard H., Orson F.M.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY158467; AAN7576.1; -;
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml; 1.
 DR SMART; SM00499; AAI; 1.
 SQ SEQUENCE 172 AA; 20114 MW; 88B991C8D8C143AB CRC64;
 Alignment Scores:
 Pred. No.: 6.04e-63 Length: 172
 Score: 815.00 Matches: 155
 Percent Similarity: 92.31% Conservative: 1
 Best Local Similarity: 91.72% Mismatches: 1
 Query Match: 64.43% Indels: 12
 DB: 2 Gaps: 1
 US-10-728-323-2 (1-717) x Q8GV20 (1-172)
 QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTGCTGCCCAAGCATCTGCGAGGCG 61
 Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaAraGln 23
 QY 62 CAGTGGGNACTCCAAGGAGACAGATGCCAGGCGAGCTCGAGGGCGGACCTGAGG 121
 Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluAraAlaLeuAraG 43
 QY 122 CCTGCGGACCAATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGAACGGAC 181
 Db 44 ProCysGluGlnHisLeuMetGlnIleGlnAraGAspGlyAraGAsp 63
 QY 182 CCTACAGCCCTAGTCAG-----GATCCG 205
 Db 64 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 83
 QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
 Db 84 TyrSerProSerProTyrAspArgGlyAlaGlySerGlnHisGlnGluAraGys 103
 QY 266 TGCATAGCTGAACGAGTTTGAGAACCAACAAAGGTGCGATGTCGAGGCGATTGCAACAG 325
 Db 104 CysAsnGluLeuAsnGluPheGluAsnGlnAraGysMetCysGluAlaLeuGlnGln 123

QY 326 ATCATCGAGAACCAAGAGAGCATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385
 Db 124 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlyAraGlnGlnGlnPheGlyAraG 143
 QY 386 GAGCTCAGGAACCTTGCTCTCAACAGTCGCGCTTAGGGCCACCAGCGTTTGCAGCTTGGAC 445
 Db 144 GluLeuAraGAsnLeuProGlnGlnCysGlyLeuAraGAlaProGlnAraGysAspLeuGlu 163
 QY 446 GTCGAAAGTGGCGGCGAGACAGATAC 472
 Db 164 ValGluSerGlyGlyAraGAspArgTyr 172
 RESULT 5
 Q6PSU2 PRELIMINARY; PRT; 187 AA.
 ID Q6PSU2; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Seed storage protein SSPI (Fragment).
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY581853; AAT00598.1; -;
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml; 1.
 DR SMART; SM00499; AAI; 1.
 FT NON TER 1
 SQ SEQUENCE 187 AA; 21786 MW; 850AF22C94983FB8 CRC64;
 Alignment Scores:
 Pred. No.: 6.07e-63 Length: 187
 Score: 815.00 Matches: 155
 Percent Similarity: 92.31% Conservative: 1
 Best Local Similarity: 91.72% Mismatches: 1
 Query Match: 64.43% Indels: 12
 DB: 2 Gaps: 1
 US-10-728-323-2 (1-717) x Q6PSU2 (1-187)
 QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTGCTGCCCAAGCATCTGCGAGGCG 61
 Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaAraGln 38
 QY 62 CAGTGGGNACTCCAAGGAGACAGATGCCAGGCGAGCTCGAGGGCGGACCTGAGG 121
 Db 39 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluAraAlaLeuAraG 58
 QY 122 CCTGCGGACCAATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGAACGGAC 181
 Db 59 ProCysGluGlnHisLeuMetGlnIleGlnAraGAspGlyAraGAsp 78
 QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
 Db 79 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 98
 QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
 Db 99 TyrSerProSerProTyrAspArgGlyAlaGlySerGlnHisGlnGluAraGys 118
 QY 266 TGCATAGCTGAACGAGTTTGAGAACCAACAAAGGTGCGATGTCGAGGCGATTGCAACAG 325
 Db 119 CysAsnGluLeuAsnGluPheGluAsnGlnAraGysMetCysGluAlaLeuGlnGln 138
 QY 326 ATCATCGAGAACCAAGAGAGCATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385

```
|||||
Db 139 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlnGlnGlnPheLysArg 158
QY 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGCGTTGCCACTTTGGAC 445
Db 159 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 178
QY 446 GTCGAAAGTGGCGGCAGACAGATAC 472
Db 179 ValGluSerGlyGlyArgAspArgTyr 187

RESULT 6
Q7Y1C0 PRELIMINARY; PRT; 169 AA.
AC Q7Y1C0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Allergen Ara h 2 isoform (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Becker W.-M., Suhr M., Lindner B., Wicklein D., Lepp U.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY117434; AA078596.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 169 AA; 19768 MW; 0B14A7ED911F34EF CRC64;

Alignment Scores:
Pred. No.: 2,49e-62 Length: 169
Score: 808.00 Matches: 154
Percent Similarity: 91.72% Conservative: 1
Best Local Similarity: 91.12% Mismatches: 2
Query Match: 63.87% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q7Y1C0 (1-169)
QY 2 CTCACCATACTAGTACCTCGCCCTTTCTCTCGCTGCCACCGCATCTCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaProAlaPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGAACTCCAAGGACACAGAAGATGCCAGAGCCAGCTCAGAGGGCGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGCAGCAACATCTCATGCAAGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 60
QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 80
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
Db 81 TyrSerProSerProTyrAspArgGlyAlaGlySerGlnHisGlnGluArgCys 100
QY 266 TGCATAGCTGAACGAGTTTGAGAACAACTGAGGATTCGAGGCAATTCGAACAG 325
Db 101 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 120
QY 326 ATCATGGAGAACAGCGGATAGTTGCGAGGGAGGCAACAGGCAACAGTTCAAGAGG 385
Db 121 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlnGlnGlnPheLysArg 140
```

```
|||||
QY 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGCGTTGCCACTTTGGAC 445
Db 141 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 160
QY 446 GTCGAAAGTGGCGGCAGACAGATAC 472
Db 161 ValGluSerGlyGlyArgAspArgTyr 169

RESULT 7
Q84TU1 PRELIMINARY; PRT; 166 AA.
AC Q84TU1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Seed storage protein (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Paik-Ro O.G., Seib J.C., Smith R.L.;
RL Theor. Appl. Genet. 104:236-240(2002).
DR EMBL; AF366560; AAO61750.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 166 AA; 19321 MW; BD04F7F26CE7B437 CRC64;

Alignment Scores:
Pred. No.: 6.31e-61 Length: 166
Score: 792.00 Matches: 151
Percent Similarity: 92.12% Conservative: 1
Best Local Similarity: 91.52% Mismatches: 12
Query Match: 62.61% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q84TU1 (1-166)
QY 2 CTCACCATACTAGTACCTCGCCCTTTCTCTCGCTGCCACCGCATCTCGAGGCAG 61
Db 2 LeuThrIleLeuValAlaProAlaPheLeuLeuAlaAlaHisAlaSerAlaArgGln 21
QY 62 CAGTGGAACTCCAAGGACACAGAAGATGCCAGAGCCAGCTCAGAGGGCGAACCTGAGG 121
Db 22 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 41
QY 122 CCTCGCAGCAACATCTCATGCAAGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
Db 42 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 61
QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
Db 62 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 81
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
Db 82 TyrSerProSerProTyrAspArgGlyAlaGlySerGlnHisGlnGluArgCys 101
QY 266 TGCATAGCTGAACGAGTTTGAGAACAACTGAGGATTCGAGGCAATTCGAACAG 325
Db 102 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 121
QY 326 ATCATGGAGAACAGCGGATAGTTGCGAGGGAGGCAACAGGCAACAGTTCAAGAGG 385
```



```
AC Q9SQG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen AraH6 (Fragment).
GN Name=Ara h 6;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
  albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274(1999).
DR ENBL; AF092846; AAD56337.1; -.
DR InterPro: IPR003612; AAI.
DR Pfam; PF00234; TRYD_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER
SQ SEQUENCE 129 AA; 14929 MW; 4048D3418B9BE40 CRC64;

Alignment Scores:
Pred. No.: 17e-25 Length: 129
Score: 388.50 Matches: 80
Percent Similarity: 66.43% Conservative: 15
Best Local Similarity: 55.94% Mismatches: 31
Query Match: 30.71% Indels: 17
DB: 2 Gaps: 4

US-10-728-323-2 (1-717) x Q9SQG5 (1-129)
QY 41 GCCCAGCATCTGCG---AGGCAGCAGTGGGAACCTCCAGGAGAC---AGAAGATGCCAG 94
Db 1 AlaHisAlaSerAlaMetArgArgGluArgGlyArgGlnGlyAspSerSerCysGlu 20
QY 95 AGCCAGCTCAGAGGGGGAACCTGAGGCCCTCGGAGCAACATCTCATGAGAAATCCAA 154
Db 21 ArgGlnValAspGlyValAsnLeuLysProCysGluGlnHisIleMetGlnArgIleMet 40
QY 155 CGTGACGAGGATTCATATGATGAACGGGACCGGTACAGCCCTAGTCAGGATCGTACAGCCCT 214
Db 41 GlyGluGlnGluGlnTyrAsp----- 47
QY 215 AGTCCATATGATCGGAGGCGCTGGATCTCTCAGCACCAGAGAGGTTGTGCAATGAG 274
Db 48 ---SerTyrAsnPheGlySerThrArgSerSerAspGlnGlnArgCysAspGlu 66
QY 275 CTGAACGAGTTTGAGAACACCAAGGTGATGTGCGAGGCATTCGACAGATCATGGAG 334
Db 67 LeuAsnGluMetGluAsnThrGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGlu 86
QY 335 AACGACGATAGTTGTCAGGGGAGCAACAGGACCAACAGTTCAAGGGGAGCTCAGG 394
Db 87 AsnGlnCysAspGlyLeuGlnAspArgGlnMetValGlnHisPheLysArgGluLeuMet 106
QY 395 AACTTGCTCAACAGTCGCGCCCTTAGGGCACCAACAGCGTTGCGACCTTGACGTCGAAAGT 454
Db 107 AsnLeuProGlnGlnCysAsnPheGlyAlaProGlnArgCysAspLeuAspVal---Ser 125
QY 455 GCGCGAGA 463
Db 126 GlyGlyArg 128

RESULT 11
Q9SQH1 PRELIMINARY; PRT; 160 AA.
ID Q9SQH1;
AC Q9SQH1;
```

```
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen.
GN Name=Ara h 7;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
  albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274(1999).
DR ENBL; AF091737; AAD56719.1; -.
DR InterPro: IPR003612; AAI.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE68808D4C CRC64;

Alignment Scores:
Pred. No.: 1.02e-17 Length: 160
Score: 300.00 Matches: 72
Percent Similarity: 61.82% Conservative: 30
Best Local Similarity: 43.64% Mismatches: 33
Query Match: 23.72% Indels: 31
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q9SQH1 (1-160)
QY 2 CTCACCATACTAGTAGCCCTC-----GCCCTTTCTCTCTCTCGCTGCC----- 43
Db 5 LeuSerIleLeuValAlaLeuLeuGlyAlaLeuLeuValAlaSerAlaThrArgTrp 24
QY 44 -----CACGCATCTCGAGGCGAGCAGTGGGAACCTC-----CAAGGAGACAGAGA 88
Db 25 AspProAspArgGlySerArgGlySerArgTrpAspAlaProSerArgGlyAspAspGln 44
QY 89 TGCCAGAGCCAGCTCGAGAGCGCGAACCTGAGGCCCTCGAGCAACATCTCATGCGAGAAG 148
Db 45 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHisMetArgArg 64
QY 149 ATCCACAGTCAGGAGTTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCGGTAC 208
Db 65 ValGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 76
QY 209 AGCCCTAGTCCATATGATCGGAGAGCGCTGGATCTCTCTCAG----- 250
Db 77 -----ProTyrSerArgArgGlySerArgGlyArgGlnProGlyGluSerAspGlu 93
QY 251 ---CACCAAGAGAGGTGTTGCAATGAGTGAACGAGTTTGAGAACACCAAGGTGCGATG 307
Db 94 AsnGlnGluGlnArgCysCysAsnGluLeuAsnArgPheGlnAsnAsnGlnArgCysMet 113
QY 308 TCGAGGAGCATTCACAGATCATGAGAACCCAGAGCGGATAG-----GTTGAGGGAGGCA 363
Db 114 CysGlnAlaLeuGlnGlnIleLeuGlnAsnGlnSerPheTrpValProAlaGlyGlnGlu 133
QY 364 ACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTAGGGC 423
Db 134 ProValAlaSerAspGlyGlyAlaGlnGluLeuAlaProGluLeuArgValGlnVal 153
QY 424 ACCACAGCGTTCGGA 438
Db 154 ThrLysProLeuArg 158

RESULT 12
Q647G8 PRELIMINARY; PRT; 158 AA.
ID Q647G8
```

```
AC Q647G8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 2S protein 2.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Aeschys.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722691; AAU21496.1; -.
SQ SEQUENCE 158 AA; 18455 MW; 32AD4D4061544948 CRC64;

Alignment Scores:
Pred. No.: 1.25e-17 Length: 158
Score: 299.00 Matches: 75
Percent Similarity: 63.12% Conservatives: 26
Best Local Similarity: 46.88% Mismatches: 39
Query Match: 23.64% Indels: 21
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q647G8 (1-158)
QY 2 CTCACCATAGTAGAGCCCTC-----GCCCTTTTCTCCTCGCTGCC----- 43
DB 4 LeuSerIleLeuValAlaLeuLeuGlyAlaLeuValValAlaSerAlaThrArgTyr 23
QY 44 -----CACGCATCTCGAGCAGCAGTGGAACTC-----CAAGGAGCAGAAGA 88
DB 24 AspProAspArgGlySerArgGlyLeuArgTyrAlaProSerArgGlyAspGln 43
QY 89 TGCAGAGCCAGCTCGAGAGGGCGAAGCTGAGGCCCTCGAGCAACTCTCATGCGAAG 148
DB 44 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHisIleArgGln 63
QY 149 ATCCAACTGACGAGGATTCATATGAAACGGGACCGCTAGCCCTAGTACGATCCGTAC 208
DB 64 ValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
QY 209 AGCCCT-----AGTCCATATGATCGAGAGGGCGCTCGATCTCTCAGCACCAGAGG 262
DB 78 SerGlnArgGlySerArgGlyArgProGlyGluSerAspGluAspGlnGlnArg 97
QY 263 TGTGCAATGAGCTGACGAGTTTGAGAACACAAAGGTGTCATGTCGAGGCGATTCGAA 322
DB 98 CysCysAsnGluLeuAsnArgPheGlnAsnAsnGlnArgCysMetCysGlnAlaLeuGln 117
QY 323 CAGATCATGAGAACACGAGATAGTTG---CAGGGGAGGCAACAGGACCACTT- 378
DB 118 GlnIleLeuGlnAsnGlnSerPheArgPheGlnGlnAspArgSerGlnLeuHisGlnAsn 137
QY 379 CAAGAGGAGCTCAGGAACCTTCCTCAACAGTGGCGCTTAGGGCACCACAGCGTTCGGA 438
DB 138 GlyGluGlyAlaGlnGluLeuAlaProGluLeuArgValGlnValThrLysProLeuArg 157

RESULT 13
Q99235 PRELIMINARY; PRT; 153 AA.
AC Q99235;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conglutin delta.
OS Lupinus angustifolius (Narrow-leaved blue lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.

OX NCBI_TaxID=3871;
RN [1]
RP SEQUENCE FROM N.A.
RA Blagrove R.J., Johnson E.D.;
RT "Biosynthesis, cDNA and amino acid sequences of a precursor of
conglutin delta, a sulphur-rich protein from Lupinus angustifolius.";
RL Plant Mol. Biol. 15: 879-893 (1990).
DR EMBL; X53523; CAA37598.1; -.
DR PIR; S12404; A33090.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 153 AA; 17785 MW; A45B7025353F0769 CRC64;

Alignment Scores:
Pred. No.: 8.7e-16 Length: 153
Score: 278.00 Matches: 62
Percent Similarity: 55.00% Conservatives: 26
Best Local Similarity: 38.75% Mismatches: 46
Query Match: 21.98% Indels: 26
DB: 2 Gaps: 5

US-10-728-323-2 (1-717) x Q99235 (1-153)
QY 2 CTCACCATAGTAGAGCCCTC-----GCCCTTTTCTCCTCGCTGCCCCACGATTCGAGG 58
DB 4 LeuThrIleLeuIleAlaLeuValAlaLeuValValHisThrSerAlaPhe 23
QY 59 CACAGTGGGAATCCACAGGACAGACAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
DB 24 Gln-----SerSerLysGlnSerCysLeuArgGlnGlnGlnValAsnLeu 39
QY 119 AGCCCTCGCAGCAACATCTCATGAGAGATCAACGTCAGCAGGATTCATATGAACGG 178
DB 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGlnGlnGluAsp 59
QY 179 GACCCGTACAGC-----CCTAGTCAGGATCCG 205
DB 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerGlnGlu 79
QY 206 TACAGCCCTAGTCCATATGATCGAGAGGGCGCTGATCTCTCAGCACCAGAGAGTGT 265
DB 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCAATGAGCTGAACGAGTTTGAGAACAAACAAAGGTGATGTCGAGGCAATTCACACAG 325
DB 91 CysGluGlnLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGGAGAACACGAGCGATAGTTTCAGGGGAGGCAACAGAGGCAACAGTTCAAGAG 385
DB 110 IleTyrGluSerGlnSerGluGlnCysGluGlySerGlnGlnGlnGlnGlnGlnGln 129
QY 386 GACCTCAGGAACCTTCCTCAACAGTGGCGCTTAGGGCACCACAGCGTTCGAGTTCGAC 445
DB 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValAsn 149

RESULT 14
2SS_SOYBN STANDARD; PRT; 158 AA.
AC P19594;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 2S albumin precursor (G2S-1) [Contains: 2S albumin small chain
DE (Aspartic acid-rich peptide); 2S albumin large chain (8 kDa
DE methionine-rich protein) (8 kDa MRP)].
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
```

```

OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv, Hodgson 78; TISSUE=Cotyledon;
RA Galvez A.F., Revilleza M.J.R., de Lumen B.O.;
RT "A novel methionine-rich protein from soybean cotyledon: cloning and
characterization of cDNA."
RL (er) Plant Gene Register PGR97-103.
RN [2]
RP SEQUENCE OF 22-64.
RC TISSUE=seed;
RX MEDLINE=87280104; PubMed=3611081;
RA Odani S., Koide T., Ono T.;
RT "Amino acid sequence of a soybean (Glycine max) seed polypeptide
having a poly(L-aspartic acid) structure."
RL J. Biol. Chem. 262:10502-10505 (1987).
RN [3]
RP SEQUENCE OF 82-96.
RA Revilleza M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;
RT "An 8 kDa methionine-rich protein (MRP) from soybean (Glycine max)
cotyledon: identification, purification and N-terminal sequence."
RL J. Agric. Food Chem. 44:2930-2935 (1996).
CC -1- FUNCTION: This is a 2S seed storage protein.
CC -1- SUBUNIT: The protein consists of two chains linked by 2 disulfide
bonds.
CC -1- SIMILARITY: Belongs to the 2S seed storage albumins family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; AF005030; AAB71140.1; -.
DR PIR; T05710; T05710.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
DR Direct protein sequencing; Seed storage protein; Signal.
KW SIGNAL
FT CHAIN 1 21
FT PROPEP 22 64 2S albumin small chain.
FT CHAIN 65 81 2S albumin large chain.
FT SITE 82 158 Cell attachment site (Potential).
FT SITE 54 56 Poly-Asp.
FT DOMAIN 56 64 Poly-Glu.
FT DOMAIN 88 91 Poly-Glu.
SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;

Alignment Scores:
Pred. No.: 1,03e-12 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x 2SS_SOYBN (1-158)

QY 5 ACCATACTAGTACCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGCGACG 64
Db 5 ThrileuLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGGACGACGAAGATGCCAGAGCCAGCTCGAGAGGCGGACCTGAGGCC 124
Db 24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
QY 125 TGGGAGCAACATCTCATGCAAGAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGlnHisIleMetGlnLysIleGlnGlyArgGlyAspAspAspAsp 62
QY 179 GACCCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGCGGCT 238
```

```

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgAsnGlu 82
QY 239 GGATCCTCTCAG-----CACCAGAGAGGCTGTTCATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGGAGAACAAACCAAGGTGCATGTCCGAGGATTCGAACAGATCATGAGAACACG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGCTTCAGGGGAGGCAACAGGACAGCAACAGTTCACAGAGGAGCTCAGGAACCTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGGCGCTTAGGGCCACACAGCGCTTGGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 15
Q9ZNZ4
ID Q9ZNZ4 PRELIMINARY; PRT; 155 AA.
AC Q9ZNZ4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Napin-type 2S albumin 1 precursor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]_TaxID=3847;
RP SEQUENCE FROM N.A.
RC STRAIN=variety PIONEER 9341; TISSUE=Seed;
RA Jung R., Hastings C., Coughlan S.J., Hu W.-N.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U71194; AAD09630.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
KW Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 155 napin-type 2S albumin 1.
SQ SEQUENCE 155 AA; 17835 MW; 2C15E5E936301E3 CRC64;

Alignment Scores:
Pred. No.: 1.4e-12 Length: 155
Score: 241.50 Matches: 58
Percent Similarity: 57.14% Conservative: 30
Best Local Similarity: 37.66% Mismatches: 53
Query Match: 19.09% Indels: 13
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q9ZNZ4 (1-155)

QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGCGAC 61
Db 4 LeuThrIleLeuLeuIleAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer 21
QY 62 CAGTGGGAACCTCAAGGACGACGAAGATGCCAGAGCCAGCTCGAGAGGCGGACCTGAGG 121
Db 22 LysTrpGlnGlnHisGlnGlnGluSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41
QY 122 CCCTGGGACCACTCATGCAAGAGATCCAA-----CGTGACGAGGATTCATAT 172
Db 42 ProCysGlu---HisIleMetGlnLysIleGlnAlaGlyArgArgGlyGlyAspGlySer 60
QY 173 GAAAGGGAGCCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA 232
Db 61 AspGluAspHisIleLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 79
QY 233 GCGCTGGATCCTCTCAG-----CACCAGAGAGGCTGTTCATGAGCTGAAC 280
```

```

Db      80  LysGluGlyLysGluGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer 99
Qy      281  GAGTTTGAGAACCAACCAAGGTGCATGTGCGAGGCATTGCAACACAGATCATGAGAACCCAG 340
Db      100  GluLeuLys--SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln 118
Qy      341  AGCGATAGGTTGCAGGGGAGGCAACAGGACACAGTTCAAGAGGGAGCTCAGGAACCTTG 400
Db      119  SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
Qy      401  CCTCAACAGTGGCGCCTTAGGGCACACACAGCGTTGCGACTTG 442
Db      138  AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

```

Search completed: August 24, 2005, 10:04:45
Job time : 47.8791 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 09:03:06 ; Search time 12.417 Seconds
(without alignments)
8620.958 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcaccatactagtagcc.....taaaagatcatgttttgg 717

Scoring table: BLOSUM62

 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat 23082005 124356 29238/app_query.fasta_1.4757
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1_46 @runat 23082005 124356 29238 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	66.6	157	4	US-09-106-872A-22
2	842	66.6	157	4	US-09-191-593-10
3	842	66.6	157	4	US-09-106-872A-2
4	827	65.4	157	4	US-09-191-593-21
5	243	19.2	158	2	US-08-618-911-4
6	243	19.2	158	3	US-08-938-675A-2
7	243	19.2	158	4	US-09-531-727-2
8	241.5	19.1	155	2	US-08-618-911-2
9	229	18.1	158	2	US-08-618-911-6
10	156	12.3	28	4	US-09-191-593-53
11	156	12.3	158	2	US-08-670-186-4
12	151	11.9	158	2	US-08-670-186-6
					Sequence 22, Appl
					Sequence 10, Appl
					Sequence 2, Appl
					Sequence 21, Appl
					Sequence 4, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 53, Appl
					Sequence 4, Appl
					Sequence 6, Appl

13	139	11.0	155	2	US-08-670-186-2	Sequence 2, Appl
14	139	11.0	184	1	US-08-453-924-3	Sequence 3, Appl
15	138	10.9	28	4	US-09-191-593-58	Sequence 58, Appl
16	136.5	10.8	164	4	US-09-627-536-13	Sequence 13, Appl
17	118	9.3	30	4	US-09-106-872A-16	Sequence 16, Appl
18	118	9.3	30	4	US-09-191-593-57	Sequence 57, Appl
19	106.5	8.4	174	4	US-09-645-593-7	Sequence 7, Appl
20	106.5	8.4	1088	4	US-09-233-857-13	Sequence 13, Appl
21	105.5	8.3	1099	3	US-09-442-100-2	Sequence 2, Appl
22	105.5	8.3	1099	4	US-08-939-106-2	Sequence 2, Appl
23	105.5	8.3	1099	4	US-09-442-102-2	Sequence 2, Appl
24	102	8.1	316	3	US-08-098-327B-31	Sequence 31, Appl
25	102	8.1	316	3	US-08-462-625-31	Sequence 31, Appl
26	100	7.9	17	4	US-09-191-593-55	Sequence 55, Appl
27	100	7.9	17	4	US-09-191-593-61	Sequence 61, Appl
28	100	7.9	17	4	US-09-191-593-62	Sequence 62, Appl
29	100	7.8	224	4	US-09-270-767-44901	Sequence 44901, A
30	99.5	7.9	148	4	US-09-083-852A-4	Sequence 4, Appl
31	99.5	7.9	148	4	US-09-083-852A-6	Sequence 6, Appl
32	99.5	7.9	148	4	US-09-489-674B-6	Sequence 6, Appl
33	97	7.7	185	4	US-09-252-991A-23490	Sequence 23490, A
34	95	7.5	32	1	US-08-158-704-1	Sequence 1, Appl
35	95	7.5	32	2	US-08-610-424B-1	Sequence 1, Appl
36	95	7.5	32	4	US-09-336-463-1	Sequence 1, Appl
37	95	7.5	32	4	US-09-191-593-1	Sequence 1, Appl
38	95	7.5	152	4	US-09-252-991A-19284	Sequence 19284, A
39	94	7.4	43	3	US-09-303-814-1	Sequence 1, Appl
40	94	7.4	500	4	US-09-265-630-13	Sequence 13, Appl
41	93.5	7.4	415	4	US-09-252-991A-18252	Sequence 18252, A
42	93	7.4	537	4	US-09-252-991A-27024	Sequence 27024, A
43	92.5	7.3	788	2	US-08-918-914-4	Sequence 4, Appl
44	91.5	7.2	1898	1	US-08-056-200-94	Sequence 94, Appl
45	91.5	7.2	1898	2	US-08-800-644-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1
US-09-106-872A-22
; Sequence 22, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannion, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-22

Alignment Scores:
Pred. No.: 8.73e-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-09-106-872A-22 (1-157)

```
QY 2 CTCACACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCACTCTCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGAACTCAAGAGACAGAAAGATGCGAGGCGAGCTCAGAGGGGGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGGAGCAACATCTCATGCAAGAGATCCACGTCGACGAGGATTCATATGAACGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCATATCATCGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCTCTCAGCAACAGAGAGGTTTGCATGATGAGCTGAAGTTTGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAenGlnArg 100
QY 302 TGCATGTGGAGGCATTGCAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCTCAACAGTCGCGCCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheIleArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGGTTGCGACTTGAGCTGCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 2
US-09-191-593-10
; Sequence 10, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
```

```
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; STRANDEDNESS: not applicable
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; DESCRIPTION: identified as Ara h II cDNA clone
; DESCRIPTION: derived amino acid sequence
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: No. 6835824 applicable
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE: seed
; TISSUE TYPE: seed cDNA
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLS: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-10
Alignment Scores:
Pred. No.: 8,73e-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-10 (1-157)
QY 2 CTCACACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCACTCTCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGAACTCAAGAGACAGAAAGATGCGAGGCGAGCTCAGAGGGGGAACCTGAGG 121
```


MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS 6.2
SOFTWARE: Wordperfect 6.0C
CURRENT APPLICATION DATA:
FILING DATE: 13 NOVEMBER 1998
FILING DATE: 04 MARCH 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/717,933
FILING DATE: 23 SEPTEMBER 1996
APPLICATION NUMBER: US 07/998,377
FILING DATE: 30 DECEMBER 1992
APPLICATION NUMBER: US 08/158,704
FILING DATE: 29 NOVEMBER 1993
APPLICATION NUMBER: US 60/009,455
FILING DATE: 29 DECEMBER 1995
APPLICATION NUMBER: US 08/610,424
FILING DATE: 04 MARCH 1996
ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
STANDARDNESS: No. 6835824 applicable
TOPOLOGY: unknown
MOLECULE TYPE: glycoprotein
STRAIN: Florunner
INDIVIDUAL ISOLATE: Ara h II
DEVELOPMENTAL STAGE:
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE:
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY: final check
LOCATION: 4..480
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-21
Alignment Scores:
Pred. No.: 4,19e-84 Length: 157
Score: 827.00 Matches: 155
Percent Similarity: 98.73% Conservative: 0
Best Local Similarity: 98.73% Mismatches: 2
Query Match: 65.38% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-21 (1-157)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCCCTGCTGCCCGCATCTCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTCGAGCAACATCTCATGCGAAGATGCCAACGTGACGAGGATTTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 61 ProSerTyrProSerGlnAspProTyrSerProTyrSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGCGATTGCAACAGATCATGAGAACCAAGAGCGATAGTTGCGAGGCGAG 361
Db 101 CysMetCysGluAlaLeuGlnGlnMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTTCCCTCAACAGTGGCGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuaspValGluSerGlyArgAspArgTyr 157
RESULT 5
US-08-618-911-4
Sequence 4, Application US/08618911
Patent No. 5850016
GENERAL INFORMATION:
APPLICANT: Jung, Rudolf
APPLICANT: Hastings, Craig
APPLICANT: Coughlan, Sean
APPLICANT: Hu, David
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
TITLE OF INVENTION: SEEDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: USA
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:


```

QY 173 GAACGGACCGTAGACCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGAGA 232
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 AspGluAspHisIleLeuIleArgThrMetPro---GlyArgIleAsnTyrIleArgLys 79
QY 233 GGCCTGAGTCTCTCAG-----CACCAAGAGAGGTGTTCATAGAGCTGAAC 280
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 LysGluGlyLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 99
QY 281 GAGTTTGAGAACAAACAGGTGATGTCGAGGCATGTCACACAGATCATGGAGAACAG 340
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 GluLeuLys---SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln 118
QY 341 AGCATAGTTCAGGGGAGGACAGAGCAACAGTTCAGAGGAGCTCAGAACTTG 400
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 SerGluGlnLeuGluGlyLys---GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
QY 401 CCTCAACAGTCGCGCCCTTAGGACCAACAGCGTTCGACATTG 442
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

```

RESULT 9

```

US-08-618-911-6
; Sequence 6, Application US/08618911
; Patent No. 5850016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; TITLE OF INVENTION: SEEDS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-618-911-6

```

Alignment Scores:

```

Pred. No.: 4.69e-17 Length: 158
Score: 229.00 Matches: 54
Percent Similarity: 55.84% Conservative: 32
Best Local Similarity: 35.06% Mismatches: 56
Query Match: 18.10% Indels: 12
DB: 2 Gaps: 6

```

US-10-728-323-2 (1-717) x US-08-618-911-6 (1-158)

```

QY 5 ACCATACTAGTACCCCTTTTCTCCTCGCTGCCACGCATCTGCGAGGCAGCAG 64
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5 ThrIleLeuLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAATCTCAAGAGAGAGAGATCCAGAGCCAGCTCGAGAGGGCGAACTCTGAGGCC 124
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuLysGlyValAsnLeuThrPro 42
QY 125 TCGAGCAACATCTCATGCAAGAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCT 238
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgLysLysGlu 82
QY 239 GATCTCTCTCAGCACCAAGAG-----AGGTGTTGCAATGAGCTGAAC 280
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 GlyLysAspGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 102
QY 281 GAGTTTGAGAACAAACAAAGTTCATGTCGAGGCATGTCACACAGATCATGGAGAACAG 340
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
103 GluLeuLys---SerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGTTGTCAGGGAGGCAACAGGAGCAACAGTTCAGAGGAGGCTCAGAACTTG 400
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 SerGluGluLeuGlu---GluLysGluAsnLysLysMetGluLysGluLeuMetAsnLeu 140
QY 401 CCTCAACAGTCGCGCCCTTAGGACCAACAGCGTTCGACATTG 442
Db   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 AlaThrMetCysArgPheGlyProMetIleGlyCysAspLeu 154

```

RESULT 10

```

US-09-191-593-53
; Sequence 53, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R

```

/ REGISTRATION NUMBER: 32,604
/ REFERENCE/DOCKET NUMBER: ARK00895601B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (501) 582-9111
/ TELEFAX: (501) 521-4931
/ TELEX: NO. 6835824 applicable
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: NO. 6835824 applicable
/ TOPOLOGY: unknown
/ MOLECULE TYPE: glycoprotein
/ DESCRIPTION: identified as derived N-terminal
/ HYPOTHETICAL: No
/ ANTI-SENSE: NO. 6835824 applicable
/ FRAGMENT TYPE: N-terminal fragment
/ ORIGINAL SOURCE:
/ ORGANISM: Arachis hypogaea
/ STRAIN: Florunner
/ INDIVIDUAL ISOLATE: P38
/ DEVELOPMENTAL STAGE:
/ HAPLOTYPE: NO. 6835824 applicable
/ TISSUE TYPE:
/ CELL TYPE: NO. 6835824 applicable
/ CELL LINE: NO. 6835824 applicable
/ ORGANELLE: NO. 6835824 applicable
/ IMMEDIATE SOURCE:
/ LIBRARY:
/ CLONE: P38
/ POSITION IN GENOME:
/ CHROMOSOME/SEGMENT: NO. 6835824 applicable
/ MAP POSITION: NO. 6835824 applicable
/ UNITS: NO. 6835824 applicable
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD: By agreement with
/ IDENTIFICATION METHOD: protein information and established
/ IDENTIFICATION METHOD: consensus sequence
/ OTHER INFORMATION: Seed storage protein and
/ OTHER INFORMATION: allergen
/ PUBLICATION INFORMATION:
/ AUTHORS:
/ TITLE:
/ JOURNAL:
/ VOLUME:
/ ISSUE:
/ PAGES:
/ DATE:
/ DOCUMENT NUMBER:
/ FILING DATE:
/ PUBLICATION DATE:
/ RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-53

Alignment Scores:
Pred. No.: 3.78e-09 Length: 28
Score: 156.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.33% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-09-191-593-53 (1-28)

QY 59 CAGCAGTGGGAATCCAAAGAGACAGAAATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
Db 1 GlnGlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeu 20
QY 119 AGGCCCTCGCAGCAACATCTCATG 142
|||||

Db 21 ArgProCysGluGlnHisLeuMet 28
RESULT 11
US-08-670-186-4
/ Sequence 4, Application US/08670186
/ Patent No. 5859343
/ GENERAL INFORMATION:
/ APPLICANT: SUN, SAMUEL S.M.
/ APPLICANT: XIONG, LIWEN
/ APPLICANT: HU, ZHONG
/ APPLICANT: CHEN, HANG
/ TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: MORRISON & FOERSTER
/ STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
/ CITY: WASHINGTON
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20006-1888
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/670,186
/ FILING DATE: 21-JUN-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MURASHIGE, KATE H.
/ REGISTRATION NUMBER: 29,959
/ REFERENCE/DOCKET NUMBER: 23461-20007.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 887-1500
/ TELEFAX: (202) 822-0168
/ TELEX: 90-4030 MRSNFOERSWSH
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 158 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-670-186-4

Alignment Scores:
Pred. No.: 7.17e-09 Length: 158
Score: 156.00 Matches: 46
Percent Similarity: 47.56% Conservative: 32
Best Local Similarity: 28.05% Mismatches: 52
Query Match: 12.33% Indels: 34
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x US-08-670-186-4 (1-158)

QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCCGTCGCCAGCATCTCGC---AGG 58
Db 5 IleLeuLeuLeuThrThrLeuAlaLeuPheValLeuLeuAlaAsnAlaSerIleTyArg 24
QY 59 CAGCAGTGGGAATCC---CAAGGAGACAGAAGATGCCAGAC 97
Db 25 ThrThrValGluLeuAspGluLeuAspGluAsnGlnProLeuCysArgArg 44
QY 98 CAGCTCGACAGG---GCGAACCTGTGAGCCCTCGAGCAACATCTCATGCAGAAATCCAA 154
Db 45 GlnPheGlnGlnHisGlnHisLeuArgAlaCysGlnArgTyIleArgArgAlaGln 64
QY 155 CGTGCAC-----GAGGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTAC 208
Db 65 ArgGlyGlyLeuValAspGluLeuGluLeuGluAspValGluGluAsnGluAsp----- 82
QY 209 AGCCCTAGTCCATATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGG----- 262
|||||


```
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-670-186-2

Alignment Scores:
Pred. No.: 5,73e-07 Length: 155
Score: 139.00 Matches: 45
Percent Similarity: 45.51% Conservative: 26
Best Local Similarity: 28.85% Mismatches: 47
Query Match: 10.99% Indels: 38
DB: 2 Gaps: 9

US-10-728-323-2 (1-717) x US-08-670-186-2 (1-155)
QY 8 ATACTAGTACCTCGCCCTTTTCTCTCGCTGCCACGCACTCGAGG----- 58
Db 7 LeuPheAlaThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleGlnThrVal 26
QY 59 -----CAGCAGTGGGAACCTCCAAGGAGACAGAAGATGC 91
Db 27 ValGluValAspGluGluAspAsnGlnLeuTip-----ArgCys 40
QY 92 CAGAGCCAG---CTCGAGAGGGCGAACCTGAGGCCCTGCGAGCAACATCTCATGCGAAG 148
Db 41 GlnArgGlnPheLeuGlnHisGlnArgLeuArgAlaCysGlnArgPheIleHisArgArg 60
QY 149 ATCCAA-----CGTGACGAGGATTTCATATGAACGGACCCGTACAGCCCTAGTCAGGAT 202
Db 61 AlaGlnPheGlyGlnProAspGluLeuGluAspGluValGluAspAsnAspAsp 80
QY 203 CGGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGGATCTCTCAGCACCAAGAGAGG 262
Db 81 GluAsnGlnPro-----ArgArgProAla-----LeuArgGln 91
QY 263 TGTTCGATGAGTGAACGAGTTTGAGAACACCAAGGTGCGATGTGC----- 310
Db 92 CysCysAsnGlnLeuArgGlnValAsp-----ArgProCysValCysProValLeuArg 109
QY 311 GAGCGATTGCAACAGACATCATGTGAGAACCCAGAGCGATAGTTGCGAGGGAGGCAACAGGAG 370
Db 110 GlnAlaGlnGlnValLeuGlnArgGlnIle-----IleGlnGlyProGlnGlnLeu 127
QY 371 CAACAGTTCAAGAGGAGCTCAGAACTTCCTCAACAGTGGCGCTT 418
Db 128 ArgArgLeuPheAspAlaAlaArgAsnLeuProAsnIleCysAsnIle 143

RESULT 14
US-08-453-924-3
/ Sequence 3, Application US/08453924
/ Patent No. 5608152
/ GENERAL INFORMATION:
/ APPLICANT: Kridl, Jean C.
/ TITLE OF INVENTION: Seed-Specific Transcriptional Regulation
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESS:
/ STREET: 2882 Sand Hill Road, Ste. 280
/ CITY: Menlo Park
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 94025
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/453,924
/ FILING DATE: 30-MAY-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/742,834
```

```
/ FILING DATE: 08-AUGUST-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/550,804
/ FILING DATE: 09-JULY-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/147,781
/ FILING DATE: 25-JANUARY-1988
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/078,538
/ FILING DATE: 28-JULY-1987
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 06/891,529
/ FILING DATE: 31-JULY-1986
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rae-Venter, Barbara
/ REGISTRATION NUMBER: 32,750
/ REFERENCE/DOCKET NUMBER: CGNE-037/04US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 926-6200
/ TELEFAX: (415) 854-3713
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 184 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-453-924-3

Alignment Scores:
Pred. No.: 6,11e-07 Length: 184
Score: 139.00 Matches: 45
Percent Similarity: 41.81% Conservative: 29
Best Local Similarity: 25.42% Mismatches: 57
Query Match: 10.99% Indels: 46
DB: 1 Gaps: 8

US-10-728-323-2 (1-717) x US-08-453-924-3 (1-184)
QY 8 ATACTAGTACCTCGCCCTTTTCTCTCGCTGCCACGCACTCTCGC---AGGCAGCAG 64
Db 8 ValSerAlaThrLeuAlaLeuPheLeuLeuThrAsnAlaSerValTyrArgThrVal 27
QY 65 TGGGAATCTCCAAAGGAGAC-----AGAGATGC 91
Db 28 ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys 47
QY 92 CAGAGCCAGCTCGAGAGGCG---AACCTGAGGCCCTGCGAGCAACATCTCATGCGAAG 148
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHisLysGln 67
QY 149 ATCCAAACGTGAC-----GAGGATTTCATATGAACGG 178
Db 68 AlaMetGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAsp 87
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
Db 88 AspValGluAsnGlnGlnGlnGlyProGlnGlnArgProProLeu----- 102
QY 239 GGATCCTCTCAGCACCAGAGAGAGGTGTCATGAGTGAACGAGTTTCAGAACCAACCA 298
Db 103 -----LeuGlnGlnCysCysAsnGluLeuHisGlnGluGlu-----Pro 115
QY 299 AGGTGCATGTGCGAGGCATTG-----CAACAGATCATGGAG 334
Db 116 LeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGlnGlnIleArgGln 135
QY 335 AACCCAGAGCATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGG----- 385
Db 136 GlnGlnGlyGlnGlnMetGlnGlnMetGlnGlnMetGlnGlnValIleSerArgIleTyrGln 155
QY 386 GAGCTCAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCCACCACAGCGTTGC 436
Db 156 ThrAlaThrHisLeuProArgAlaCysAsnIleArgGlnValSerIleCys 172
```

RESULT 15
US-09-191-593-58
; Sequence 58, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M.
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6835824 applicable
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; DESCRIPTION: identified as 17.5 kD N-terminal
; DESCRIPTION: sequence of Ara h 2 p38 (gene sequence 20..47).
; HYPOTHETICAL: No
; ANTI-SENSE: No. 6835824 applicable
; FRAGMENT TYPE: amino terminus
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE: P38
; DEVELOPMENTAL STAGE:
; HAPLOTYPE: No. 6835824 applicable
; TISSUE TYPE:
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLE: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY: P38
; CLONE: P38

; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-58
Alignment Scores:
Pred. No.: 3.95e-07 Length: 28
Score: 138.00 Matches: 26
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 2
Query Match: 10.91% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-58 (1-28)
QY 59 CAGCAGTGGAACTCCAAAGGAGACAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
Db 1 GlnGlnTrpGluLeuGlnGlyAspArgArgGlnSerGlnLeuGluArgAlaAsnLeu 20
QY 119 AGGCCCTGCGAGCAACATCTCATG 142
Db 21 ArgProCysGluGlnHisLysMet 28
Search completed: August 24, 2005, 10:09:12
Job time : 15.417 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 24, 2005, 10:05:17 ; Search time 46.2283 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US10728323/runat_23082005_124357_29323/app_query.fasta_1.4757
-DB=Published Applications AA -QFMT=fastan -SUFFI=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10728323@cgn 1 1 221 @runat 23082005 124357_29323
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	842	66.6	157	14	US-10-228-806-4 Sequence 4, Appl
2	842	66.6	157	15	US-10-100-303A-63 Sequence 63, Appl
3	838.5	66.3	207	17	US-10-958-324-2 Sequence 2, Appl
4	835	66.0	160	17	US-10-899-551-4 Sequence 4, Appl
5	819	64.7	156	15	US-10-245-871-53 Sequence 53, Appl
6	819	64.7	156	15	US-10-253-286-53 Sequence 53, Appl
7	771	60.9	166	15	US-10-100-303A-82 Sequence 82, Appl
8	760	60.1	167	17	US-10-899-551-56 Sequence 56, Appl
9	699	55.3	166	9	US-09-731-221-77 Sequence 77, Appl
10	463	36.6	83	15	US-10-100-303A-81 Sequence 81, Appl
11	243	19.2	158	15	US-10-302-633-2 Sequence 2, Appl
12	243	19.2	158	15	US-10-424-599-169225 Sequence 169225,
13	241.5	19.1	155	15	US-10-424-599-264372 Sequence 264372,
14	177.5	14.0	168	14	US-10-165-289A-2 Sequence 2, Appl
15	149.5	11.8	167	16	US-10-382-086-29 Sequence 29, Appl
16	144	11.4	169	14	US-10-165-289A-4 Sequence 4, Appl
17	140.5	11.1	165	16	US-10-425-115-190971 Sequence 190971,
18	139	11.0	184	11	US-09-782-130-3 Sequence 3, Appl
19	139	11.0	184	14	US-10-100-121-48 Sequence 48, Appl
20	129.5	10.2	150	15	US-10-442-174A-1 Sequence 1, Appl
21	127.5	10.1	151	15	US-10-424-599-167336 Sequence 167336,
22	127	10.0	122	14	US-10-228-806-67 Sequence 67, Appl
23	126	10.0	102	15	US-10-424-599-219338 Sequence 219338,
24	124	9.8	148	16	US-10-425-115-219350 Sequence 219350,
25	119	9.4	20	9	US-09-731-221-66 Sequence 66, Appl
26	118	9.2	261	15	US-10-291-265-715 Sequence 715, App
27	117	9.2	20	9	US-09-731-221-63 Sequence 63, Appl
28	116	9.2	20	9	US-09-731-221-60 Sequence 60, Appl
29	115	9.1	20	9	US-09-731-221-61 Sequence 61, Appl
30	113	8.9	20	9	US-09-731-221-65 Sequence 65, Appl
31	112.5	8.9	101	15	US-10-424-599-275916 Sequence 275916,
32	112	8.9	20	9	US-09-731-221-54 Sequence 54, Appl
33	112	8.9	20	9	US-09-731-221-59 Sequence 59, Appl
34	112	8.9	20	9	US-09-731-221-62 Sequence 62, Appl
35	112	8.9	20	9	US-09-731-221-64 Sequence 64, Appl
36	111	8.8	20	9	US-09-731-221-57 Sequence 57, Appl
37	111	8.8	20	9	US-09-731-221-74 Sequence 74, Appl
38	111	8.8	20	9	US-09-731-221-75 Sequence 75, Appl
39	111	8.8	20	9	US-09-731-221-76 Sequence 76, Appl
40	110	8.7	25	15	US-10-245-871-60 Sequence 60, Appl
41	110	8.7	25	15	US-10-253-286-60 Sequence 55, Appl
42	109	8.6	20	15	US-10-245-871-55 Sequence 55, Appl
43	109	8.6	20	15	US-10-253-286-55 Sequence 244, App
44	109	8.6	149	18	US-10-481-032A-244 Sequence 53, Appl
45	108	8.5	20	9	US-09-731-221-53

ALIGNMENTS

RESULT 1
US-10-228-806-4
; Sequence 4, Application US/10228806
; Publication NO. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-4

Alignment Scores:


```
QY 122 CCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTCGAGAGCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGGCAATTCGAACAGATCATGCAGAACCCAGAGCGATAGTTCGAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAGAGCTTCGCACTTGGAGCTGAAAGTGGCGGCGAGACAGATCTAAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgArgProArgGileProPro 163
QY 482 ATCTCAAAAAAGAAAGAAAGAAAGAAATAGCTTATATATAGC---TATTATCTA 538
Db 164 IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183
QY 539 TGGTTATGTTAGTT 553
Db 184 AlaMetCysLeuLeu 188

RESULT 4
US-10-899-551-4
; Sequence 4, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 160
; TYPE: PRT
; ORGANISM: species Arachis hypogaea
US-10-899-551-4

Alignment Scores:
Pred. No.: 6,31e-79 Length: 160
Score: 835.00 Matches: 156
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DB: 17 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-4 (1-160)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGACG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGAACTCCAAAGAGACAGAAAGATGCCAGAGCGAGTTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTCGAGAGCAACCAAGG 301
```

```
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTCGAGAGCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGGCAATTCGAACAGATCATGCAGAACCCAGAGCGATAGTTCGAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAGAGCTTCGCACTTGGAGCTGAAAGTGGCGGCGAGACAGACAGA 469
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArg 159

RESULT 5
US-10-245-871-53
; Sequence 53, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-53

Alignment Scores:
Pred. No.: 3,06e-77 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-2 (1-717) x US-10-245-871-53 (1-156)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGACG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGAACTCCAAAGAGACAGAAAGATGCCAGAGCGAGTTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTCGAGAGCAACCAAGG 301
```

Db 84 SerSerGlnHisGlnGlnArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGCATTCGACAGATCATCGAGAACACAGAGCGATAGGTTCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGATTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGATGCGGCCCTTAGG 421
Db 124 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 6
US-10-253-286-53
; Sequence 53, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PROT
; ORGANISM: Arachis hypogaea
US-10-253-286-53
Alignment Scores:
Pred. No.: 3,06e-77 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
Gaps: 15
US-10-728-323-2 (1-717) x US-10-253-286-53 (1-156)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTGCGAGGCAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCAAGAGACAGAAGATGCCAGAGCCAGCTCAGAGGGCGGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCTGGAGCAACATCTCATGACAGAGATCCACGTCGACGAGATTTCATATGAACGGGAC 181
Db 44 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGlnAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGGTTTGCATGACCTGAACAGTTTTCAGAACCAACGAAGG 301
Db 84 SerSerGlnHisGlnGlnArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTGCGAGCATTCGACAGATCATCGAGAACACAGAGCGATAGGTTCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGATGCGGCCCTTAGG 421

Db 124 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 7
US-10-100-303A-82
; Sequence 82, Application US/10100303A
; Publication No. US20030202960A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 166
; TYPE: PROT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-82
Alignment Scores:
Pred. No.: 3,61e-72 Length: 166
Score: 771.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
Gaps: 15
US-10-728-323-2 (1-717) x US-10-100-303A-82 (1-166)
QY 50 TCTGCGAGCGACAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCGAGCTCGAGAGG 109
Db 15 SerAlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArg 34
QY 110 GCGAACCCTGAGCCCTGCGAGCAACATCTCATGCAAGAGATCCAAAGTACGAGGATTTCA 169
Db 35 AlaAsnLeuArgProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer 54
QY 170 TATGAACGGGACCCGTACAGCCCTAGTCCGATCCGTACAGCCCTAGTCCATATGATCGG 229
Db 55 TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 74
QY 230 AGAGCGCTGGATCTCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTGAG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu 94
QY 290 ACAACCAAGGTGCATGTGCGAGGCATTCGACAGATTCGACAGATCATGAGAACCGAGCTAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
QY 350 TTGCGAGGAGGCAACAGGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTGCTCAACAG 409
Db 115 LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln 134
QY 410 TCGCGCCTTAGGGCCACCACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGCAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 154
QY 470 TAC 472
Db 155 Tyr 155
RESULT 8
US-10-899-551-56
; Sequence 56, Application US/10899551
; Publication No. US2005006394A1
; GENERAL INFORMATION:


```
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 167
; TYPE: PRT
; ORGANISM: species Arachis hypogaea
US-10-899-551-56

Alignment Scores:
Pred. No.: 5,23e-71 Length: 167
Score: 760.00 Matches: 139
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.08% Indels: 0
DB: 17 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-56 (1-167)
Qy 53 GCGAGGAGCAGTGGGAACCTCCAGGAGACAGAGAAGATGCCAGAGCCAGCTCGAGAGGCG 112
Db 17 AlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAla 36
Qy 113 AACCTGAGGCCCTCGAGCAACATCTCATGACAGAAGATCCAACTGACGAGGATTCATAT 172
Db 37 AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyr 56
Qy 173 GAACGGACCGGTACAGCCCTAGTCAGATCCGACAGCCCTAGTCATATGATCGAGA 232
Db 57 GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 76
Qy 233 GCGCTGGATCCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAC 292
Db 77 GlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsn 96
Qy 293 AACCAAGGTGCATGTGCGAGGCATTCACAGATCATGAGAACACGAGCGATAGTTG 352
Db 97 AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeu 116
Qy 353 CAGGGAGGCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCTTGCCTCAACAGTGC 412
Db 117 GlnGlyArgGlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCys 136
Qy 413 GGCCTTAGGGCACCACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCAGAGACAGA 469
Db 137 GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 155

RESULT 9
US-09-731-221-77
; Sequence 77, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
```

```
; OTHER INFORMATION: Hypogaea
US-09-731-221-77

Alignment Scores:
Pred. No.: 1,42e-64 Length: 166
Score: 699.00 Matches: 131
Percent Similarity: 92.91% Conservatives: 0
Best Local Similarity: 92.91% Mismatches: 10
Query Match: 55.26% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x US-09-731-221-77 (1-166)
Qy 50 TCTGCGAGGAGCAGTGGGAACCTCCAGGAGACAGAGAAGATGCCAGAGCCAGCTCGAGAG 109
Db 15 SerAlaArgGlnGlnAlaGluLeuGlnGlyAspArgCysGlnSerGlnLeuAlaArg 34
Qy 110 GCGAACCTGAGGCCCTCGAGCAACATCTCATGACAGAAGATCCAACTGACGAGGATTC 169
Db 35 AlaAsnLeuArgAlaCysGluAlaHisLeuMetGlnLysIleGlnAlaAspGluAspSer 54
Qy 170 TATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG 229
Db 55 TyrGluArgAlaProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArg 74
Qy 230 AGAGGCGCTGGATCCTCTCAGCAACGAGAGAGGTGTTGCAATGAGCTGAACGAGTTGAG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGlu 94
Qy 290 AACCAACCAAGGTGCATGTGCGAGGCATTCGCAACAGATCATGAGAACACGAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
Qy 350 TTCAGCGGGAGGCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCTTGCCTCAACAG 409
Db 115 LeuGlnGlyAlaGlnGlnGlnGlnPheLysArgGluAlaArgAsnLeuProGlnGln 134
Qy 410 TGGCGCTTAGGGCACACAGCCTTGGGACTTGGAGCTTCGAAAGTGGCGGCAGAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArg 154
Qy 470 TAC 472
Db 155 Tyr 155

RESULT 10
US-10-100-303A-81
; Sequence 81, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-81

Alignment Scores:
Pred. No.: 9,37e-40 Length: 83
Score: 463.00 Matches: 83
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 36.60% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-2 (1-717) x US-10-100-303A-81 (1-83)
```

Qy	68	GA	ACTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTTGAGGCCCTGC	127
Db	1	GL	uLeuGLnGLyAspArgArgCysGLnSerGLnLeuGLuArgAlaAenLeuArgProCys	20
Qy	128	GAGCAACATCTCATGCAGAGATCCAACGTCACGAGGATTCATATGAACGGACCCGTAC	187	
Db	21	GLuGLnHisLeuMetGLnLysIleGLnArgAspGLuAspSerTyGLuA-GAspProTy	40	
Qy	188	AGCCCTAGTCAGGATCCGTACAGCCCTAGTTCATATGATCGAGAGCGCTGGATCCTCT	247	
Db	41	SerProSerGLnAspProTy-SerProSerProTyAspArgArgGLyAlaGLySerSer	60	
Qy	248	CAGCACCAAGAGAGGTGTTCGAATCAGCTGACGAGTTCAGAGTTCAGAAACCAAGAGTGCATG	307	
Db	61	GLnHisGLnGLuArgCysCysAsnGLuAenGLuPheGLuAenAsnGLnArgCysMet	80	
Qy	308	TGCAGAGCA	316	
Db	81	CysGLuAla	83	
		RESULT 11		
		US-10-302-633-2		
		Sequence 2, Application US/10302633		
		Publication No. US20030229038A1		
		GENERAL INFORMATION:		
		APPLICANT: de Lumen, Benito O.		
		Galvez, Alfredo F.		
		TITLE OF INVENTION: Lunasin Peptides		
		NUMBER OF SEQUENCES: 3		
		CORRESPONDENCE ADDRESS:		
		ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP		
		STREET: 75 DENISE DRIVE		
		CITY: HILLSBOROUGH		
		STATE: CALIFORNIA		
		COUNTRY: USA		
		ZIP: 94010		
		COMPUTER READABLE FORM:		
		MEDIUM TYPE: Floppy disk		
		COMPUTER: IBM PC compatible		
		OPERATING SYSTEM: PC-DOS/MS-DOS		
		SOFTWARE: Patentin Release #1.0, Version #1.30		
		CURRENT APPLICATION DATA:		
		APPLICATION NUMBER: US/10302,633		
		FILING DATE: 22-No. US20030229038A1-2002		
		CLASSIFICATION: <Unknown>		
		PRIOR APPLICATION DATA:		
		APPLICATION NUMBER: US/09/531,727		
		FILING DATE: 21-Mar-2000		
		APPLICATION NUMBER: 08/938,675		
		FILING DATE: <Unknown>		
		ATTORNEY/AGENT INFORMATION:		
		NAME: OSMAN, RICHARD A		
		REGISTRATION NUMBER: 36,627		
		REFERENCE/DOCKET NUMBER: B98-003		
		TELECOMMUNICATION INFORMATION:		
		TELEPHONE: (650) 343-4341		
		TELEFAX: (650) 343-4342		
		INFORMATION FOR SEQ ID NO: 2:		
		SEQUENCE CHARACTERISTICS:		
		LENGTH: 158 amino acids		
		TYPE: amino acid		
		STRANDEDNESS: single		
		TOPOLOGY: linear		
		MOLECULE TYPE: peptide		
		SEQUENCE DESCRIPTION: SEQ ID NO: 2:		
		US-10-302-633-2		
		Alignment Scores:		
		Pred. No.:	1,828-16	Length: 158
		Score:	243.00	Matches: 56
		Percent Similarity:	56.49%	Conservative: 31
		Best Local Similarity:	36.36%	Mismatches: 55

Qy	65	TGGGA	ACTC	AGGAG	ACAG	AGATG	CCAG	AGGC	AC	CTCG	AGAG	GGCG	AA	CC	TGAG	GC	CC	124
		:::																
Db	24	TrpGln	HisGln	Asp	---	SerCys	ArgLys	GlnLeu	GlnGly	ValAsn	LeuThr	Pro	42					
		:::																
Qy	125	TGGCAG	CAAC	ATCT	CAT	CGAC	GAAG	ATCCAA	-----	CGTG	ACGAGG	ATTCAT	TATGA	ACGG	178			
Db	43	CysGlu	LysHis	IleMet	GlnLys	IleGln	GlyArg	GlyAsp	AspAsp	AspAsp	Asp	62						
Qy	179	GACCCG	TACAG	CCCT	TAGT	CAGG	ATCCG	TACAG	CCCT	TAGT	CCAT	TATGAT	CGAG	AGGCGCT	238			
Db	63	AspAsp	AsnHis	IleLeu	ArgThr	MetArg	GlyArg	Ile	AsnTyr	IleArg	ArgAsn	Glu	82					
Qy	239	GGATCT	CTCT	CG	-----	-----	-----	-----	-----	CAC	ACAG	AGAG	GTGTT	CCAT	TATGAT	CTGAAC	280	
Db	83	GlyLys	AspGlu	AspGlu	AspGlu	GluGlu	GluGly	HisMet	GlnLys	CysCys	ThrGlu	MetSer	102					
Qy	281	GAGTTT	GACAA	CAAC	CAAG	TGTCAT	GTGC	AGGCAT	TTGCA	ACAGAT	CATG	TGAG	AA	CCAG	340			
Db	103	GluLeu	---	ArgSer	Pro	LysCys	GlnCys	LysAla	LeuGln	LysIle	MetGlu	AsnGln	121					
Qy	341	AGCGAT	GAGTT	CGGGG	AGGCA	ACAGG	ACCA	CAGTT	TCAG	AGGGG	AGTCT	CAGAA	ACTTG	400				
Db	122	SerGlu	GluLeu	Glu	---	GluLys	GlnLys	LysLys	MetGlu	LysGlu	LeuIle	AsnLeu	140					
Qy	401	CCTCA	ACAGT	TGGG	CGCTT	AGGC	ACCA	CAGCGT	TGCG	ACTTG	442							
Db	141	AlaThr	MetCys	ArgPhe	GlyPro	MetIle	GlnCys	AspLeu	154									

```

RESULT 13
US-10-424-599-264372
; Sequence 264372, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264372
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80749C.1.pep
US-10-424-599-264372

```

Alignment Scores:	
Pred. No.:	2.6e-16
Score:	241.50
Percent Similarity:	57.14%
Best Local Similarity:	37.66%
Query Match:	19.09%
DB:	15
Length:	58
Matches:	58
Conservative:	30
Mismatches:	53
Indels:	13
Gaps:	7

US-10-728-323-2 (1-717) x US-10-424-599-264372 (1-155)

QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTTCCTCCTGCGCCACGCATCTCGCAGGCAG	61
		::: ::: ::: :::	
Db	4	LeuThrIleLeuLeuIlealaLeuLeuPheIIe-----AlaHisThrCysCysAlaSer	21
		::: ::: ::: :::	
QY	62	CAGTGGGAACCTCCAAGGACACAAGAAGTCCAGACGCCAGCTCGAGGCGCGCAACTCAGG	121
		::: ::: ::: ::: :::	
Db	22	LysTrpGlnGlnHisGlnGlnGlusuCysArgGluGlnLeuLysGlyIleAsnLeuAasn	41
		: ::: ::: ::: :::	
QY	122	CCGTGGGACCAATCTCATGCAGAAAGATCCAA-----CGTGACGAGGATTTCATAT	177
		::: ::: ::: :::	
Db	42	ProCysGlu---HisIleMetGluLysIleGlnAlaGlyArgArgGlyGluAspGlySer	60
		::: ::: ::: :::	

QY	173	GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA	232
Db			
Db	61	AspGluAspHisIleLeuIleArgThrMetPro--GlyArgIleAsnTrpIleAsGlyLys	79
QY	233	GGCGTGTGATCCTCTCAG-----CACCAAGAGAGCTGTTGCAATGAGCTGAAC	280
Db	80	LysGluGlyLysGluGluGluGluGluGluGluHisMetGlnLysCysSerGluMetSer	99
QY	281	GAGTTTGAGAACAAACCAAGGTGCATGTGCGAGGCGATTGCAACAGATCATGAGAAACCAAG	340
Db	100	GluLeuLys--SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln	118
QY	341	AGCGATGGTTGCGGGGAGGCAACAGCAGGACCAACAGTTCACAGAGGAGCTCAGGAACCTTG	400
Db	119	SerGluGlnLeuGluGlyLys---GluLysLysGlnMetGluArgGluLeuMetAsnLeu	137
QY	401	CCTCAACAGTCGGCGCTTAGGCGACCAACAGCGGTTGCGACTTG	442
Db	138	AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu	151

```

RESULT 14
US-10-165-289A-2
; Sequence 2, Application US/10165289A
; Publication No. US20030159174A1
; GENERAL INFORMATION:
; APPLICANT: QIU, Xiao
; APPLICANT: TRUSKA, Martin
; APPLICANT: HU, Zhiyuan
; TITLE OF INVENTION: Flax (Linum usitatissimum L.)
; TITLE OF INVENTION: Seed-Specific Promoters
; FILE REFERENCE: BNZ-005
; CURRENT APPLICATION NUMBER: US/10/165,289A
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/295823
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Linum usitatissimum
US-10-165-289A-2

```

Alignment Scores:		
Pred. No.:	1.5e-09	Length:
Score:	177.50	Matches:
Percent Similarity:	45.71%	Conservative:
Best Local Similarity:	30.86%	Mismatches:
Query Match:	14.03%	Indels:
DB:	14	Gaps:
		9

US-10-728-323-2 (1-717) x US-10-165-289A-2 (1-168)

Qy	2	CTCACATACTAGTACGCTCGCC-----CTTTTCCTCTCGTGGCCACGCATCT	52
Db	4	LeuMetSerLeuAlaValAlaAlaThrAlaPheLeuPheLeuIleValValAspAlaSer	23
Qy	53	GCAGGACGACGATGGAACTCCAGGACAC-----	82
Db	24	ValArgThrThrValIleIleAspGluAspThrAsnGlnGlyArgGlyGlyGlnGlyGly	43
Qy	83	-----AGAAGATGCCAGGCAGCTC---GAGAGGCGCAACTGAGGCCCTGTC	121
Db	44	GlnGlyGlnGlnGlnCysGluLysGlnIleGlnGlnAspTyrLeuArgSerCys	63
Qy	128	GAGCAACATCTCATCAGACAGATCCAACTGACGAGGATTCATATGAACGGACCCGTAC	187
Db	64	GlnGlnPheLeuTrpGluLysValGlnLysGlyArgSerTyr-----	78
Qy	188	AGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA-----	241
Db	79	-----TTTTrAsnGlnGlyValArgGlyGlnGlyGly	87

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 4512.08 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcgcagcaaccgaggagaa.....ctccgaggctgtgcttaa 1524

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1524	100.0	1524	6	BD175911
2	1524	100.0	1524	8	AF093541
3	1509.6	99.1	1524	6	AX155334
4	1391.2	91.3	1853	6	AX148741
5	1391.2	91.3	1853	8	AF086821
6	1373.2	90.1	1590	8	AF125192
7	1329.4	87.2	1901	8	AY722686
8	1328	87.1	1934	8	AY722685
9	1321.6	86.7	1886	8	AY439332
10	854.2	56.0	1533	8	AF510854
11	834.6	54.8	3825	8	AF510854
12	694.8	45.6	1786	8	AY722687
13	519.2	34.1	657	8	AF487543
14	517.4	34.0	1446	6	BD175913
15	517.4	34.0	1446	6	AR202578
16	517.4	34.0	1446	6	AR486465
17	517.4	34.0	1446	6	AX839958
18	517.4	34.0	1446	6	AX840483
19	515.8	33.8	1696	8	AB030494

20	513	33.7	1488	6	BD175911	BD175911
21	513	33.7	1488	6	AR202576	AR202576
22	513	33.7	1488	6	AR486463	AR486463
23	513	33.7	1488	6	AX839950	AX839950
24	513	33.7	1488	6	AX840475	AX840475
25	513	33.7	1550	8	AB113349	AB113349
26	513	33.7	1743	6	E32414	E32414
27	513	33.7	1743	6	AR343131	AR343131
28	513	33.7	1743	8	SOYGLYBSU	M36866 Soybean gly
29	505.6	33.2	1708	8	SOYGA2B1A	D00216 Glycine max
30	505	33.1	1746	6	E02463	E02463 cDNA encodi
31	505	33.1	1746	6	AX252313	AX252313
32	505	33.1	1746	8	GMGLY1A	X02985 Soybean mRN
33	504	33.1	1458	6	BD175912	BD175912
34	504	33.1	1458	6	AR202577	AR202577
35	504	33.1	1458	6	AR486464	AR486464
36	504	33.1	1458	6	AX839952	AX839952
37	504	33.1	1458	6	AX840477	AX840477
38	504	33.1	1639	8	AB113350	AB113350
39	502.4	33.0	1712	6	E02462	E02462 cDNA encodi
40	502.4	33.0	1712	8	GMGLY1	X02806 Glycine max
41	502.2	33.0	1710	8	AB030495	AB030495
42	477.4	31.3	1729	8	PSA132614	AJ132614 Pisum sat
43	431.8	28.3	1664	8	VSLEGUMA	Z32835 V.sativa mR
44	426.2	28.0	1666	8	VFLEA2	X55014 Vicia faba
45	421.8	27.7	1850	8	VNPROLEGA	Z46803 V.narbonens

ALIGNMENTS

RESULT 1
BD107900
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD107900
Methods and reagents for decreasing allergic reactions.
BD107900
BD107900.1 GI:23202718
JP 2002501748-A/3.
unidentified
unclassified.
1 (bases 1 to 1524)
Sosin, H., Banon, G.A., Jr, W.A.B. and Samphson, H.A.
Methods and reagents for decreasing allergic reactions
Patent: JP 2002501748-A 3 22-JAN-2002;
UNIVERSITY OF ARKANSAS, MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY
UNIVERSITY OF NEW YORK, HOWARD SOSIN
OS Arachis L. (Peanut)
PN JP 2002501748-A/3
PD 22-JAN-2002
PF 29-JAN-1999 JP 2000529437
PR 31-JAN-1998 US 60/073283,13-FEB-1998 US 60/074590 PR
13-FEB-1998 US 60/074624,13-FEB-1998 US 60/074633 PR
27-AUG-1998 US 09/141220
PI HAWADO SOSIN, GARY A BANON WESLEY A BIRX JR, HYU A SAMPTHON PC
C12N15/09,A01H5/00,A01K67/027,A61K39/35,C07K14/37,C07K14/415, PC
C07K14/435,
PC C07K14/47,C12N1/15,C12N1/19,C12N1/21,C12N5/00,C12N5/00,C12N5/00
PC 00
CC Methods and reagents for decreasing allergic reactions FH
Key source Location/Qualifiers
FT
FT
Location/Qualifiers
1. 1524
/organism="Arachis L. (Peanut)";
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1524; DB 6; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1  CGCAGCAACCGGAGGAGAACGGGTGCCAGTTTCAGCGCCTCAATCGCAGAGACCTGAC  60
Db      1  CGCAGCAACCGGAGGAGAACGGGTGCCAGTTTCAGCGCCTCAATCGCAGAGACCTGAC  60
QY     61  AATCGCATGGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACGAGAGTTC  120
Db     61  AATCGCATGGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACGAGAGTTC  120
QY    121  GAATGCGCGCGCTCGCCCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGGCT  180
Db    121  GAATGCGCGCGCTCGCCCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGGCT  180
QY    181  TTCTACTCCCAATGCTCCCAAGGAGATCTTCATCAGCAAGGAAGGGGATCTTTGGGTTG  240
Db    181  TTCTACTCCCAATGCTCCCAAGGAGATCTTCATCAGCAAGGAAGGGGATCTTTGGGTTG  240
QY    241  ATATTCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAGAGTGTGATCTCAG  300
Db    241  ATATTCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAGAGTGTGATCTCAG  300
QY    301  TCCCAAGACCCAAAGAGCTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT  360
Db    301  TCCCAAGACCCAAAGAGCTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT  360
QY    361  CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCAITGCACTTCCACCGGTGTGCT  420
Db    361  CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCAITGCACTTCCACCGGTGTGCT  420
QY    421  TTCTGGCTCTAACAGACACGACACTGATGTTGTTGCTGTTTCTTACTGACACCAAC  480
Db    421  TTCTGGCTCTAACAGACACGACACTGATGTTGTTGCTGTTTCTTACTGACACCAAC  480
QY    481  AACAAACGACACCGAGTTGATCAGTTCCCGCAGGAGATTCAATTTGCTGGGAACCGAG  540
Db    481  AACAAACGACACCGAGTTGATCAGTTCCCGCAGGAGATTCAATTTGCTGGGAACCGAG  540
QY    541  CAAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT  600
Db    541  CAAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT  600
QY    601  AGCCCATACGCCCGCAAGTACGCTAGACAAGAGAGCGGTGAATTTAGCCCTCAGGA  660
Db    601  AGCCCATACGCCCGCAAGTACGCTAGACAAGAGAGCGGTGAATTTAGCCCTCAGGA  660
QY    661  CAGCAGACGCGCAGAGAACGAGCAGACGAAGAAGAAACCAAGGTGGAACATCTTC  720
Db    661  CAGCAGACGCGCAGAGAACGAGCAGACGAAGAAGAAACCAAGGTGGAACATCTTC  720
QY    721  AGCGGCTTCACGCGGAGTTCTTGGAAACAGCCCTCCAGGTTGACGACAGACAGATAGTG  780
Db    721  AGCGGCTTCACGCGGAGTTCTTGGAAACAGCCCTCCAGGTTGACGACAGACAGATAGTG  780
QY    781  CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGGGGA  840
Db    781  CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGGGGA  840
QY    841  GGCTCAGAACTTTGAGCCCGAGATAGAAAGAGACGTCGCGACGAAGAGGAATACGAT  900
Db    841  GGCTCAGAACTTTGAGCCCGAGATAGAAAGAGACGTCGCGACGAAGAGGAATACGAT  900
QY    901  GAAGATGAATATGAATACGATGAAGAGGATAGAGGCGGTGGCAGGGGAAGCAGAGCAG  960
Db    901  GAAGATGAATATGAATACGATGAAGAGGATAGAGGCGGTGGCAGGGGAAGCAGAGCAG  960
QY    961  GGGAAATGGTATGAAGAGAGGATCTGCACCGCAAGTCTTAAAGAAACATTGGTAGAAAC  1020
Db    961  GGGAAATGGTATGAAGAGAGGATCTGCACCGCAAGTCTTAAAGAAACATTGGTAGAAAC  1020
QY   1021  AGATCCCTGACATCTTACAAACCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC  1080
Db   1021  AGATCCCTGACATCTTACAAACCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC  1080
```

```
QY      1081  CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTTACAGGAATGCA  1140
Db      1081  CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTTACAGGAATGCA  1140
QY     1141  TTGTTTGTGCTCACTACAAACCAACGACACACAGCATCATATATCGATTGAGGGACGG  1200
Db     1141  TTGTTTGTGCTCACTACAAACCAACGACACACAGCATCATATATCGATTGAGGGACGG  1200
QY    1201  GGTCAAGTCAAGTGTGTCAGACCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG  1260
Db    1201  GGTCAAGTCAAGTGTGTCAGACCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG  1260
QY    1261  GGTCAAGTGTGTCGCCACAGAACTTCGCCGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG  1320
Db    1261  GGTCAAGTGTGTCGCCACAGAACTTCGCCGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG  1320
QY    1321  TTCGAATAGTGGCATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA  1380
Db    1321  TTCGAATAGTGGCATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA  1380
QY    1381  AACTCCGTCTATAGATAACCTCGCGAGGAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG  1440
Db    1381  AACTCCGTCTATAGATAACCTCGCGAGGAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG  1440
QY    1441  GAGCAGGCAAGCAGCTTAAGAAACAAACACCCCTTCAAGTCTTCCACCGTCTCAG  1500
Db    1441  GAGCAGGCAAGCAGCTTAAGAAACAAACACCCCTTCAAGTCTTCCACCGTCTCAG  1500
QY    1501  CAGTCTCCGAGGCTGTGGCTTAA  1524
Db    1501  CAGTCTCCGAGGCTGTGGCTTAA  1524

RESULT 2
AF093541      1524 bp      mRNA      linear      PLN 31-OCT-2001
LOCUS      Arachis hypogaea glycinin (Arah3) mRNA, partial cdb.
DEFINITION      AF093541
ACCESSION      AF093541
VERSION      AF093541.1  GI:3703106
KEYWORDS      .
SOURCE      Arachis hypogaea (peanut)
ORGANISM      Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE      1 (bases 1 to 1524)
AUTHORS      Rabjohn, P., Helm, E.M., Stanley, J.S., West, C.M., Sampson, H.A.,
Burks, A.W. and Bannon, G.A.
Molecular cloning and epitope analysis of the peanut allergen Ara h
3
J. Clin. Invest. 103 (4), 535-542 (1999)
JOURNAL
MEDLINE      99146968
PUBMED      10021462
REFERENCE      2 (bases 1 to 1524)
AUTHORS      Rabjohn, P., Helm, E.M., Stanley, J.S., West, C.M., Sampson, H.A.,
Burks, A.W. and Bannon, G.A.
Direct Submission
Submitted (21-SEP-1998) Pediatrics, University of Arkansas for
Medical Sciences, 1120 Marshall Street, Little Rock, AR 72202, USA
FEATURES
source
1..1524
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Florunner"
/db_xref="taxon:3818"
<1..1524
/gene="Arah3"
<1..1524
/gene="Arah3"
/feature="seed storage; peanut allergen"
/codon_start=1
/product="glycinin"
gene
CDS
```


ORIGIN		/db_xref="taxon:3818"	
Query Match		99.1%; Score 1509.6; DB 6; Length 1524;	
Best Local Similarity		99.4%; Pred. No. 0;	
Matches 1515; Conservative		0; Mismatches 9; Indels 0; Gaps 0;	
QY	1	CGGCAGCAACCGGAGGAGAACCGGTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
DB	1	CGGCAGCAACCGGAGGAGAACCGGTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
QY	61	AATCGCATTAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACGAGAGTTC	120
DB	61	AATCCCAATTGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACGAGAGTTC	120
QY	121	GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCTT	180
DB	121	GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCTT	180
QY	181	TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGATATCTTTGGGTTG	240
DB	181	TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGATATCTTTGGGTTG	240
QY	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGGCTTCAACAAAGGTCGTGATCTCAG	300
DB	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGGCTTCAACAAAGGTCGTGATCTCAG	300
QY	301	TCCCAAGACCAACAGAGGCTTCCAGGAGAGCAACCAAGCCCAACAGCAGAGATGT	360
DB	301	TCCCAAGACCAACAGAGGCTTCCAGGAGAGCAACCAAGCCCAACAGCAGAGATGT	360
QY	361	CACGAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT	420
DB	361	CACGAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT	420
QY	421	TTCTGGCTCTCAACAGCACAAGCACTGATGTTGTTGCTGTTTCTCTTACTGACACCAAC	480
DB	421	TTCTGGCTCTCAACAGCACAAGCACTGATGTTGTTGCTGTTTCTCTTACTGACACCAAC	480
QY	481	AACAACGACCAACAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGACACGGAG	540
DB	481	AACAACGACCAACAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGACACGGAG	540
QY	541	CAAGAGTTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCATAT	600
DB	541	CAAGAGTTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCATAT	600
QY	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
QY	661	CAGCAGCGCGCAGAGACGAGCAGCAAGAGAGAGAGAAACGAAGGTGGAACATCTTC	720
DB	661	CAGCAGCGCGCAGAGACGAGCAGCAAGAGAGAGAGAAACGAAGGTGGAACATCTTC	720
QY	721	AGCGGCTTCACGCGCGAGTTCTGGAACAAGCCTTCAGGTTGACACAGACAGATAGTG	780
DB	721	AGCGGCTTCACGCGCGAGTTCTGGAACAAGCCTTCAGGTTGACACAGACAGATAGTG	780
QY	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
DB	781	CAAAACCTTAAGAGCGCAGACTGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
QY	841	GGCCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGCAAGAGAGGAATACGAT	900
DB	841	GGCCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGCAAGAGAGGAATACGAT	900
QY	901	GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGGCGG	960
DB	901	GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGGCGG	960
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTTAAAAAGAACATTGGTAGAAAC	1020
DB	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTTAAAAAGAACATTGGTAGAAAC	1020
QY	1021	AGATCCCTTGACATCTCAACACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC	1080
DB	1021	AGATCCCTTGACATCTCAACACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC	1080
QY	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA	1140
DB	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA	1140
QY	1141	TTGTTTGTGCTCACTCAACACCAACGACACAGCATCATATATCGATTGAGGGACCG	1200
DB	1141	TTGTTTGTGCTCACTCAACACCAACGACACAGCATCATATATCGATTGAGGGACCG	1200
QY	1201	GCTCAGCTCAAGTCTGTGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG	1260
DB	1201	GCTCAGCTCAAGTCTGTGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG	1260
QY	1261	GCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAAC	1320
DB	1261	GCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAAC	1320
QY	1321	TTGCAATACGTGGCATTTCAAGACAGACTCAAGGCCCGACATAGCCAACTCGCGGTGAA	1380
DB	1321	TTGCAATACGTGGCATTTCAAGACAGACTCAAGGCCCGACATAGCCAACTCGCGGTGAA	1380
QY	1381	AACTCGCTCATAGATAAAGCTCGCGGAGAGGTGGTTGCAATTCATATGGCTTCAAAAG	1440
DB	1381	AACTCGCTCATAGATAAAGCTCGCGGAGAGGTGGTTGCAATTCATATGGCTTCAAAAG	1440
QY	1441	GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG	1500
DB	1441	GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG	1500
QY	1501	CAGTCTCCGAGGCTGTGGCTTAA	1524
DB	1501	CAGTCTCCGAGGCTGTGGCTTAA	1524
RESULT 4			
AX148741		1853 bp	DNA
LOCUS		Sequence 4 from Patent WO0136621.	linear
DEFINITION		AX148741	PAT 08-JUN-2001
ACCESSION		AX148741.1	
VERSION		GI:14347295	
KEYWORDS		Arachis hypogaea (peanut)	
SOURCE		Arachis hypogaea	
ORGANISM		Arachis hypogaea	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;	
		Aeschynomeneae; Arachis.	
REFERENCE		1	
AUTHORS		Dodo, H.W., Arntzen, C.J., Konan, K.N. and Viquez, O.M.	
TITLE		Down-regulation and silencing of allergen genes in transgenic	
		peanut seeds	
JOURNAL		Patent: WO 0136621-A 4 25-MAY-2001;	
		Alabama A & M University (US)	
FEATURES		Location/Qualifiers	
		source	
		1..1853	
		/organism="Arachis hypogaea"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3818"	
ORIGIN			
Query Match		91.3%; Score 1391.2; DB 6; Length 1853;	
Best Local Similarity		95.4%; Pred. No. 0;	
Matches 1454; Conservative		0; Mismatches 68; Indels 2; Gaps 2;	
QY	2	GGCAGCAACCGGAGGAGAACCGGTGCAGTTCAGCGCTCAATGGCAGAGACCTGACA	61
DB	71	GGCAGCAACCGGAGGAGAACCGGTGCAGTTCAGCGCTCAATGGCAGAGACCTGACA	130

QY 62 ATGCGATTGAATCAGAGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGAGTTGC 121
Db 131 ACCGATTTGAATCGAGAGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGAGTTGC 190
QY 122 AATCGCGCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACCGCCCTTCGTAGGCGCTT 181
Db 191 AATCGCGCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACCGCCCTTCGTAGGCGCTT 250
QY 182 TCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAATCTTTGGGTTGA 241
Db 251 TCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAATCTTTGGGTTGA 310
QY 242 TATTCCCTGGTTGCTTAGACACTATGAAGAGCTCTCACACAGAGTCTCGATCTCAGT 301
Db 311 TATTCCCTGGTTGCTTAGACACTATGAAGAGCTCTCACACAGAGTCTCGATCTCAGT 370
QY 302 CCAAAGACCAACAGAGCTCTCCAGGAGGAAGACCAAGCCCAACAGCAACAGATAGTC 361
Db 371 CCAAAGACCAACAGAGCTTTGCAAGAGAAGAACCAAGCCCAACAGCAACAGATAGTC 430
QY 362 ACAGAAAGTGCCACCGTTTCGATGAGGGTGAATCTCATTTGCGATTTCCACCGGTGTCCTT 421
Db 431 ACCAAGAGGTGCACCGTTTCAATGAGGGTGAATCTCATTTGCGATTTCCACCGGTGTCCTT 490
QY 422 TCTGCTCTCAACAGCACAGACACTGATGTTGCTGCTTCTTCTTACTGACACCAACA 481
Db 491 TCTGCTCTCAACAGCACAGACACTGATGTTGCTGCTTCTTCTTACTGACACCAACA 550
QY 482 ACAACGACAAACAGCTTGATCAGTTCCCCAGGAGATTCATTTTGGCTGGGAACACGAGC 541
Db 551 ACAACGACAAACAGCTTGATCAGTTCCCCAGGAGATTCATTTTGGCTGGGAACACGAGC 610
QY 542 AAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAGAGCAGCAAGAGCTTACCATATA 601
Db 611 AAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAGAGCAGCAAGAGCTTACCATATA 670
QY 602 GCCATACAGCCCGCAAGTCAGCTAGACAAGAGAGCGTGATTTAGCCCTCGAGGAC 661
Db 671 GCCATACAGCCCGCAAGTCAGCTAGACAAGAGAGCGTGATTTAGCCCTCGAGGAC 730
QY 662 AGCACAGCCGAGAGACGAGCAGACAAGAGAAGAAACAGAGGTGGAACATCTTCA 721
Db 731 AGCACAGCCGAGAGACGAGCAGACAAGAGAAGAAACAGAGGTGGAACATCTTCA 790
QY 722 GCGCTTCACCGCGAGTTCTTGAAACAAGCTTCCAGTTGACGACAGACAGATAGTC 781
Db 791 GCGCTTCACCGCGAGTTCTTGAAACAAGCTTCCAGTTGACGACAGACAGATTTGTC 850
QY 782 ARAACCTTAAGGCGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGAG 841
Db 851 ARAATCTGTGGGGCGAGAACGAGAGTGAAGAAGAGGAGCCATTGTGACGCTGAAGGAG 910
QY 842 GCCTCAGAACTTTGAGCCAGATAGAGAAGAGACGTGCGCAGCAAGAGAAGAGATACCATG 901
Db 911 GCCTCAGAACTTTGAGCCAGATAGAGAAGAGAGGTGCGCAGCAAGAGAAGAGATACCATG 970
QY 902 AAGATGAATATGAATACGATGAAGAGATGAAGCGCTGCGAGGGGAAGCAGAGCGCAGG 961
Db 971 AAGATGAATATGAATACGATGAAGAGATGAAGCGCTGCGAGGGGAAGCAGAGCGCAGG 1030
QY 962 GGATGCTATTGAAGAGACGATCTGCACCGCAGTGTCTAAAAGAACATTTGGTAGAACA 1021
Db 1031 GGAATGCTATTGAAGAGACGATCTGCACCGCAGTGTCTAAAAGAACATTTGGTAGAACA 1090
QY 1022 GATCCCTGACATCTACAACCCCTCA-AGCTGGTTTCACTCAAACTGCAACGATCTCAAC 1080
Db 1091 GATCCCTGACATCTACGATCTCAGCGCTGTTTCACTCAAACTGCGC-ACGATCTCAAC 1149
QY 1081 CTTCTAATATCTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTTCTCAGGAATGCA 1140
Db 1150 CTTCTAATCTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTTCTCAGGAATGCA 1209
QY 1141 TTGTTTTCGCTCACTACAACACCAACGACACAGCATCATATCATTCATTTAGGGGACGG 1200

Db 1210 TTGTTTTCGCTCACTACAACCAACGACACAGATCATATATGCTTTGAGGGACGG 1269
QY 1201 GCTCAGCTGCAAGTCTGGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1270 GCTCAGCTGCAAGTCTGGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1329
QY 1261 GGTCAAGTCTGTTGTTGGTCCACAGAACTTTCGCGTTCGCTGGAAAGTCCACAGAGGAGAAC 1320
Db 1330 GGTCAAGTCTTCTGTTGGTCCACAGAACTTTCGCGTTCGCTGGAAAGTCCACAGAGGAGAAC 1389
QY 1321 TTGCAATACGTGTCATTTCAAGACAGACTCAAGGCCACAGCATAGCCAACTTCGCGGTGAA 1380
Db 1390 TTGCAATACGTGTCATTTCAAGACAGACTCAAGGCCACAGCATAGCCAACTTTGCGGTGAA 1449
QY 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGTGTGTCAAATTCATATGCGCTCCAAAGG 1440
Db 1450 AACTCCGTCTATAGATAACCTGCGGAGGAGTGTGTCAAATTCATATGCGCTCCCAAGG 1509
QY 1441 GACGAGCAAGGAGCTTAAAGAACAAACCCCTTCAGTTCCTGTTCCACGCTCTCAG 1500
Db 1510 GACGAGCAAGGAGCTTAAAGAACAAACCCCTTCAGTTCCTGTTCCACGCTCTCAG 1569
QY 1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
Db 1570 CAGTCTCCGAGGCGTGTGGCTTAA 1593

RESULT 5
AF086821
LOCUS Arachis hypogaea glycinin (AraH4) mRNA linear PLN 29-SEP-1999
DEFINITION Arachis hypogaea glycinin (AraH4) mRNA, complete cds.
ACCESSION AF086821
VERSION AF086821.1 GI:5712198
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 1853)
Becker, W.M.
Selective cloning of peanut allergens, including profilin and 2S
albumins, by phage display technology
Int. Arch. Allergy Immunol. 119 (4), 265-274 (1999)
JOURNAL MEDLINE 99406463
PUBMED 10474031
REFERENCE 2 (bases 1 to 1853)
Kleber-Janke, T.
Direct Submission
TITLE Submitted (25-AUG-1998) Biochemische und Molekulare Allergologie,
Forschungszentrum Borstel, Parkallee 22, Borstel 23845, Germany
JOURNAL Location/Qualifiers
FEATURES
source 1..1853
/organism="Arachis hypogaea"
/mol_type="mRNA"
/strain="Virginia"
/db_xref="taxon:3818"
/dev_stage="seed"
1..1853
/genes="AraH4"
1..1593
/genes="AraH4"
/notes="seed storage protein; allergen"
/codon_start=1
/product="glycinin"
/protein_id="AAD47382.1"
/db_xref="GI:5712198"
/translation="MAKLELSPFCFLVLGASSISFRQOPEENACQFORLNAORPDN
RISEGGYLETWNPNNQPEFCAGVALSLVLRLNALRRPFYSNAPQEIFTQQGRGVFG
LIFPGCPSTYEPAQQGRYQSORPPRLQEEDSQSQDQSHQKVRHFNEDGLIAVPT

GVAFWLYNDHDTDVAVSLTDNNNDNQLDQPRFRNLAGNHEQEFLRYQQOSRORR
RSLPYSPYSHPREREFRRGQHSRRAGOEDEEGNIFSGFTPEFLEQAO
VDROQIVONLWNGNEBEGAIIVTRGSLRIILSPDTRGADBEEDVEYEHEDG
RCGRSGRGNGGTEETICTACVKKNIGNESPHIYDPQRFTQNDHDLNLLILWLGL
SAGYGNLYRNALFVPHYNTNAHSIIIALKRGRAHVQVDSNGRVYDEEIQEGHVLVVP
QNFVAGKQSENFEVFAFKTDSRPSIANFAGNSFIDNLPEEVVANSYGLPREQARQ
LKNPNKFEFVPPFOQSPRAVA"

ORIGIN

Query Match	91.3%;	Score	1391.2;	DB	8;	Length	1853;
Best Local Similarity	95.4%;	Pred. No.	0;				
Matches	1454;	Conservative	0;	Mismatches	68;	Indels	2;
Gaps	2;						

QY	2	GGCAGCAACCGGAGGAGACGCGTGCAGTTCAGGCGCTCAATGGCGCAGAGACCTTGACA	61
DB	71	GGCAGAGCCGGAGGAGATGCGTGCAGTTCAGGCGCTCAATGGCGCAGAGACCTTGACA	130
QY	62	ATCGCAATTGAATCAGAGGGCGGTTACATTGAGACTTTGGAAACCCCAACCAACGAGAGTTGC	121
DB	131	ACGCAATTGAATCGGAGGGCGGTTACATTGAGACTTTGGAAACCCCAACCAACGAGAGTTGC	190
QY	122	AATGCGCGCGGTCGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCGCTTCGTAGGCGTT	181
DB	191	AATGCGCGCGGTCGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCGCTTCGTAGGCGTT	250
QY	182	TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTTGA	241
DB	251	TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTTGA	310
QY	242	TATTCCCTGTTGCTCTAGACACTATGAGAGGCTCACACAAAGTGTCGATCTCAGT	301
DB	311	TATTCCCTGTTGCTCTAGACACTATGAGAGGCTCACACAAAGTGTCGATCTCAGT	370
QY	302	CCCAAGACCAACAGACGCTCCAGGAGAACCAAGCCCAACAGCAACGAGATAGTC	361
DB	371	CCCAAGACCAACAGACGCTTCGAAGAGAACCAAGCCCAACAGCAACGAGATAGTC	430
QY	362	ACCAGAAGTGCAACCGTTTCGATGAGGGTGATCTCATTTGCAAGTTCACACCGGTGTTGCTT	421
DB	431	ACCAGAAGTGCAACCGTTTCGATGAGGGTGATCTCATTTGCAAGTTCACACCGGTGTTGCTT	490
QY	422	TCTGGCTCTACAAACGACCAACGACACTGATGTTGCTGTTGCTCTTTATGACACCAACA	481
DB	491	TCTGGCTCTACAAACGACCAACGACACTGATGTTGCTGTTGCTCTTTATGACACCAACA	550
QY	482	ACAACGACACACGACTGATCTCCCGAGGAGATTCATTTGCTGGGACACCGAGC	541
DB	551	ACAACGACACACGACTGATCTCCCGAGGAGATTCATTTGCTGGGACACCGAGC	610
QY	542	AAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGCAAGAAAGCTTACCATATA	601
DB	611	AAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGCAAGAAAGCTTACCATATA	670
QY	602	GCCCATACAGCCCGCAAGTACGCTTAGACAGAGCGGTGAATTTAGCCCTCGAGGAC	661
DB	671	GCCCATACAGCCCGCAAGTACGCTTAGACAGAGCGGTGAATTTAGCCCTCGAGGAC	730
QY	662	AGCACAGCCGACAGACGAGGAGGACAGAAAGCAAGAAAGGAGTGGAAACATCTTCA	721
DB	731	AGCACAGCCGACAGACGAGGAGGACAGAAAGCAAGAAAGGAGTGGAAACATCTTCA	790
QY	722	GCGGCTTACGCGGAGTTCCTCGAAACAAAGCCTTCCAGGTTGACGACAGACATAGTGC	781
DB	791	GCGGCTTACGCGGAGTTCCTCGAAACAAAGCCTTCCAGGTTGACGACAGACATAGTGC	850
QY	782	AAACCTAGAGCGGACGACGAGGTGAAGAGAGGAGCCATTGTGACAGTGGGGAG	841
DB	851	AAAATCTGTGGGCGAAGACGAGGTGAAGAGAGGAGCCATTGTGACAGTGGGGAG	910
QY	842	GCCTCAGAACTTTGAGCCCAATAGAAAGAGAGCTGCCGACGAGAGAGGATACGATG	901
DB	911	GCCTCAGAACTTTGAGCCCAATAGAAAGAGAGCTGCCGACGAGAGAGGATACGATG	970

QY	902	AAGATGAATATGAATACGATGAAGAGGATAGAAAGCGTGCAGGGGAAGCAGAGCGGAG	961
DB	971	AAGATCAATATGAATACGATGAAGAGGATAGAAAGCGTGCAGGGGAAGCAGAGCGGAG	1030
QY	962	GGAAATGGTATTGAAGAGAGCATCTGCACCGCAAGTGCTTAAAAAGAAACATTGGTAGAAACA	1021
DB	1031	GGAAATGGTATTGAAGAGAGCATCTGCACCGCATGTGTAAAAAGAAACATTGGTAGAAACA	1090
QY	1022	GATCCCTCGACATCTACAAACCTCTCA-AGCTGGTGTCTCACTCAAAACCTGCCAAACGATCTCAAC	1080
DB	1091	GATCCCTCGACATCTACGATCTCTCAGCGCTGGTGTCTCACTCAAAACCTGCC-ACGATCTCAAC	1149
QY	1081	CTTCTAATACTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA	1140
DB	1150	CTTCTAATACTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA	1209
QY	1141	TTGTTTGTGCTCATCTACAAACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
DB	1210	TTGTTTGTGCTCATCTACAAACCAACGACACAGCATCATATATGAGGGGACGG	1269
QY	1201	GCTCAGTGCAGTGTGGACAGCAACGCAACAGAGTGTACGAGGAGCTTCAAGAG	1260
DB	1270	GCTCAGTGCAGTGTGGACAGCAACGCAACAGAGTGTACGAGGAGCTTCAAGAG	1329
QY	1261	GCTCAGTGTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCGAGAGCGAGAAC	1320
DB	1330	GCTCAGTGTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCGAGAGCGAGAAC	1389
QY	1321	TTGCAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCGGTGAA	1380
DB	1390	TTGCAATACGTGGCATTTCAAGACAGACTTTCAAGGCCAGCATAGCCAACTTTGCGGTGAA	1449
QY	1381	AACTTCGCTCATAGATAACCTCGCGGAGGAGTGTGTTGCAAAATTCATATGGCTTCCAAAGG	1440
DB	1450	AACTTCGCTCATAGATAACCTCGCGGAGGAGTGTGTTGCAAAATTCATATGGCTTCCAAAGG	1509
QY	1441	GAGCAGGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCCTCGTCCACCGTCTCAG	1500
DB	1510	GAGCAGGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCCTCGTCCACCGTCTCAG	1569
QY	1501	CAGTCTCCGAGGCTGTGSCITTA 1524	
DB	1570	CAGTCTCCGAGGCTGTGSCITTA 1593	

RESULT 6

AF125192

LOCUS

DEFINITION

Arachis hypogaea Gly1 (Gly1) mRNA, complete cds.

ACCESSION

AF125192

VERSION

AF125192.2

GI:9864776

KEYWORDS

Arachis hypogaea (peanut)

SOURCE

ORGANISM

Arachis hypogaea

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 1590)

Jain, A.K. and Basha, S.M.

Molecular cloning of a glycinin-type peanut seed storage protein

TITLE

Unpublished

REFERENCE

2 (bases 1 to 1590)

Jain, A.K. and Basha, S.M.

Direct Submission

TITLE

Submitted (02-FEB-1999) Plant Biotechnology, Florida A&M University, 301 South Perry Paige Building, Tallahassee, FL 32307, USA

REFERENCE

3 (bases 1 to 1590)

Jain, A.K. and Basha, S.M.

Direct Submission

TITLE

Submitted (22-AUG-2000) Plant Biotechnology, Florida A&M University, 301 South Perry Paige Building, Tallahassee, FL 32307, USA

ORGANISM

Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 1901)
CDNA clone of peanut seed storage protein gene

AUTHORS

Yan, Y., Wang, L. and Huang, S.
Unpublished

JOURNAL

2 (bases 1 to 1901)
Direct Submission
Submitted (30-JUL-2004) Department of Biology and Biotechnology,
Sun Yat-sen University, Guangzhou, Guangdong 510275, China

FEATURES

Location/Qualifiers

1..1901

/organism="Arachis hypogaea"

/mol_type="mRNA"

/db_xref="taxon:3818"

30..1643

/notes="seed storage protein"

/codon_start=1

/product="arachin Ahy-2"

/protein_id="AAU21491.1"

/db_xref="GI:52001221"

/translations="MAKLIALSVCFVLVGASSISFRQPPBENACQFORLNAQRPD
RLESEGYIETWPNQEPFECAGVALSLVLRNALRRPFYSNAPQIEFIQGRGVFG
LIFPGPSNVEEPAQGRHQSORAPREFEGDSQOQDQSHQVRRDEGDLIAP
TGVALMYNDHDQDVAVSLTDNNNDQLDQPRFNLAGNHEQBFLLYQQSRRS
LPYSPYQSDQREERFSPROGRRERAGQENEGNIFSGFTPEFLAQAFQVD
DROILNLRGENSDQGAIVTVRGLRLILSPDKRRQYRPERDEEYDEYEDYE
EERQDRRRGSRGRNGIETICTASVKINGRNSPDIVNPOAGSLKTDNLNL
ILRWLGLSREYGLNLYNALFVPHYNTNAHSIIIVLRGRAHVQVVDNGNRVYDEIQE
GHVLVVPQFAVAKSQSDNFVAFKTDSPRIANLAGENSIIIDNLPEEVVANSGL
PREQARQLKNNPFKTFVPPSQSLGAVA"

ORIGIN

Query Match 87.2%; Score 1329.4; DB 8; Length 1901;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1448; Conservative 0; Mismatches 66; Indels 39; Gaps 4;

Qy 2 GGCAGCAACCGAGGAGAACCGGTGCAGTTCAGCGCTCAATGCGCAGAGACCTTGACA 61
Db |||||
Qy 100 GGCAGCAGCCGAGGAGAAATGCGTGCAGTTCAGCGCTCAATGCGCAGAGGCGCTGACA 159
Db |||||
Qy 62 ATCGCATTGAATCAGAGGCGGTTACATTGAGACTTGGAACTTGGACCCCAACACAGAGTTG 121
Db |||||
Qy 160 ACCGCTTTGAATCGAGGCGGTTACATTGAGACTTGGAACTTGGAACTTGGAACTTGG 219
Db |||||
Qy 122 AATGCGCGCGCTCGCTCTCTCGTTAGTTCCTCGCGCGCAACGCGCTTCGTAGGCGCTT 181
Db |||||
Qy 220 AATGCGCGCGCTCGCTCTCTCGTTAGTTCCTCGCGCGCAACGCGCTTCGAGGCGCTT 279
Db |||||
Qy 182 TCTACTCAATGCTCCCGAGGAGATCTTCAATCAGCAGAGGAGGATCTTTGGTTGA 241
Db |||||
Qy 280 TCTACTCAATGCTCCCGAGGAGATCTTCAATCAGCAGAGGAGGATCTTTGGTTGA 339
Db |||||
Qy 242 TATTCCCTGCTGCTCTAGACACTATGAGAGGCTTCAACACAGGCTCGTCTGATCTCAGT 301
Db |||||
Qy 340 TATTCCCTGCTGCTCTAGACACTATGAGAGGCTTCAACACAGGCTCGTCTGATCTCAGT 399
Db |||||
Qy 302 CCCAAAGACCAACCAAGACGCTCTCAAGGAGAGACCAACCAAGGAGGAGGAGGAGGAG 358
Db |||||
Qy 400 CGCAAGAGACCAACAGAGCTTTTGAAGGAGAGACCAACCAAGGAGGAGGAGGAGGAG 459
Db |||||
Qy 359 GTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 418
Db |||||
Qy 460 GTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 519
Db |||||
Qy 419 CTCTTCTGCTCTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 478
Db |||||
Qy 520 CTCTTCTGATGTAAACGAGCAGTGAACACTGATGTTGTTGCTGTTCTCTTACTGACACCA 579
Db |||||

RESULT 8
AY722685 1934 bp mRNA linear PLN 19-SEP-2004
LOCUS Arachis hypogaea arachin Ahv-1 mRNA, complete cds.
DEFINITION Arachis hypogaea arachin Ahv-1 mRNA, complete cds.
ACCESSION AY722685
VERSION AY722685.1 GI:52001218
KEYWORDS Arachis hypogaea (peanut)
SOURCE Arachis hypogaea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 1934)
AUTHORS Yan, Y., Wang, L. and Huang, S.
TITLE cDNA clone of peanut seed storage protein gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1934)
AUTHORS Yan, Y., Wang, L. and Huang, S.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2004) Department of Biology and Biotechnology, Sun Yat-sen University, Guangzhou, Guangdong 510275, China
FEATURES Location/Qualifiers
source 1..1934
organism="Arachis hypogaea"
mol_type="mRNA"
db_xref="taxon:3818"
59..1669
note="seed storage protein"
codon_start=1
product="arachin Ahv-1"
protein_id="AAU21490.1"
db_xref="GI:52001219"
translation="MGKLLALVCFCFLVILGASSISFQQOPEENACQFRLNAQRDPN
RISSEGYIETWPNNEPFCAGVALSELVLRNLRPEYSSNAPQEIFTQQGRGVFG
LIPFCGSTTEEPAGQGRHQSPRRPFGQQDQSQQQSHQKVRHFDGDLIAVPT
GVAFMYNDHDTDVAVSLDTNNNDQLDQFPRFLAGNHEQFLRYEQOQRSSRL
PYSPYSPQTPQKDEDFSPRGHRRERAGQEENEGNI FSGFTPEFLAQAFQVDD
RQILLNQGENSDQCAI VTVRGLRILSPDRKQOQYERPDDEEYDEDEYDEDE
ERQHDRRGSRGSGNGIRETICTASFKNIGNRSPDIYNPQAGSLKTANLELLLI
RLWLGLSAEYGNLYRNALFVPHYNTNAHSIIYALRGRAHVQVVDNSGDRVDEDELOGL
HLVLPVFNFAVGSGSENFEYAFKTDTSRPSSTANLAGENSIFIDNLPPEVVAANSYGLP
REQARQDKNNPNPFKFFPPEQSLRAVA"
ORIGIN
Query Match 87.1%; Score 1328; DB 8; Length 1934;
Best Local Similarity 92.8%; Pred No. 0;
Matches 143; Conservative 0; Mismatches 75; Indels 36; Gaps 3;

Qy 2 GGCAGCAACCGGAGGAGAACGCGTGCCAGTGTTCAGCGCCTCAATGCGCAGAGACTTGACA 61
|
Db 129 GGCAGCAGCGGAGGAGAAATCGTGCAGCTTCAGCGCCTCAATGCGCAAGGCGCTGACA 188
|
Qy 62 ATGCGATTGAATCAGAGCGGGGTTCATTGAGACTTGGAACCCCAACACAGAGAGTTCG 121
|
Db 189 ACCGATTGAATCGAGCGGGGTTCATTGAGACTTGGAACCCCAACACAGAGAGTTCG 248
|
Qy 122 AATGCGCGGCGTGCCTCTCTCGCTTAGTCTTCGCGCAACGCGCCTTCGTAGGCGCTT 181
|
Db 249 AATGCGCGGCGTGCCTCTCTCGCTTAGTCTTCGCGCAACGCGCCTTCGTAGGCGCTT 308
|
Qy 182 TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTGTA 241
|
Db 309 TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTGTA 368
|
Qy 242 TATTCCCTGTTGTCCTAGACATATGAAGAGCCTCAACACAGAGTTCGTGATCTCAGT 301
|
Db 369 TATTCCCTGTTGTCCTAGACATATGAAGAGCCTCAACACAGAGTTCGTGATCTCAGT 428
|
Qy 302 CCCAAGACCAACCAAGACGCTTCCAAGGAGAGACCAAGCCACACAGACAGATAGTC 361
|
Db 429 CCCAAGACCAACCAAGACGCTTCCAAGGAGAGACCAAGCCACACAGACAGATAGTC 488
|

Db 1392 TGGCTGGAAAGTCCCGACGAGCAGAACTTTGTAATACGTGGGCATTTCAAGACAGACTCAAGGC 1451

Qy 1355 CCAGCATAGCCAACTCCCGCGTGAATACTCCGTATAGATAAACCCTGCCGAGGAGGTGG 1414

Db 1452 CCAGCATAGCCAACTAGCCCGTGAATACTCTTCATAGATACTTTCGCGGAGGAGGTGG 1511

Qy 1415 TTGCAAAATTCATATGGCTCCAAAGGAGCAGGCAAGGACGCTTAAAGAACAAACCCCT 1474

Db 1512 TTGCAAAATTCATATGGCTCCCAAGGAGCAGGCAAGGACGCTTAAAGAACAAACCCCT 1571

Qy 1475 TCAAGTTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524

Db 1572 TCAAGTTCTTCGTTCCACCGTCTGAACAGTCTCTGAGGGCTGTGGCTTAA 1621

RESULT 10

AY618460 1533 bp mRNA linear PLN 07-JUN-2004

LOCUS Arachis hypogaea glycinin mRNA, partial cds.

DEFINITION Arachis hypogaea glycinin mRNA, partial cds.

ACCESSION AY618460

VERSION Arachis hypogaea L. GI:47933674

KEYWORDS

SOURCE Arachis hypogaea (peanut)

ORGANISM Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

REFERENCE 1 (bases 1 to 1533)

AUTHORS Kang, I.-H. and Gallo-Meagher, M.

TITLE Cloning and characterization of a novel ara h 3, a major peanut (Arachis hypogaea L.) allergen gene

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 1533)

AUTHORS Kang, I.-H. and Gallo-Meagher, M.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2004) Agronomy, University of Florida, Gainesville, FL 32611, USA

FEATURES

source 1..1533

/organism="Arachis hypogaea"

/mol_type="mRNA"

/db_xref="taxon:3818"

<1..1533

/note="seed storage protein; Ara h 3-im; peanut allergen"

/codon_start=1

/product="glycinin"

/protein_id="AAT39430.1"

/db_xref="GI:47933675"

/translation="KLIALSLCFVLGASVTFRQGENECQFORLNAQRPDNRRI ESEGGYLETWPNNOBFQCAVLASRTVLARNALRPFYSNAPLELYVOOGSGVGLI PFGCPSTYEPAGRGYRQSKSRRFQVQDDPSQQQDSHQKHVRFDPGDLIAVPT GVAFMYNDDQTVLTSIDTSIHNQLOFPFRFLVAGNQEFLRYQQQSGRPH YRQISPRVDQEQNESNIFSGPAQVDFLOHAFQVDTVENLRGNEREQGAVITV KGLRLILSPDESSRSPRREEDFSRQQRKYDENRRGYNKGTETTCGAS VKNLGRSNNPDIYNPAQSLRSWNELDPLIGWLSAQHGTIYRNAMFVPHYTLNA HTIIVALNGRAHVQVDNNGRVYDELOEGHVLVDPQNFVAAKAQSENVEYLAFT DSRPSIANLAGENSIINLPEEVVANSYRLPREQARQLKNPNFKFPVFPFDQSMRE VA"

ORIGIN

Query Match 56.0%; Score 854.2; DB 8; Length 1533;

Best Local Similarity 76.1%; Pred. No. 1.8e-212;

Matches 1177; Conservative 0; Mismatches 286; Indels 102; Gaps 6;

Qy 2 GGCAGCAACCGGAGGAGAACCGTGCAGTTTCCAGCGCCTCAATGCGCAGAGACCTGACA 61

Db 65 GGCAGGGGGGGAGGAGGAATGAGTGCCAGTTCCAGCGCCTCAATGCGCAGAGACCGACA 124

Qy 62 ATCCGATTGAATCAGAGGGCGGTTACATTGAGACTTGGAACCCCAACAAACGAGGAGTTGC 121

Db 125 ACCGCAATTGAGTCTTGAAGGGCGGTTACATTGAGACTTGGAACCCCAACAAACGAGGAGTTCC 184


```
Db      1187  TTTGGCAATTGAACGGACGGGCACATGTGCAAGTGTGGGACGCAACCGGTAAACAGAGTGT 1246
Qy      1241  ACACGAGGAGCTTCAAGAGGGTCAAGTGTCTTGTGGTGCACAGAACTTCGCCGTGCGCTG 1300
Db      1247  ACCACGAGGAGCTTCAAGAGGGTCAAGTGTCTTGTGGTGCACAGAACTTCGCCGTGCGAG 1306
Qy      1301  GAAAGTCCAGAGCGAGAACTTCGAATACGTGGCAATTCAGACAGACTCAAGGCCCGCAGCA 1360
Db      1307  CAAAGGCCAGAGCGAGAACTACGAATACCTTGGCGTTCAAGACAGACTCAAGGCCCGCAGCA 1366
Qy      1361  TAGCCAACTCGCGGTGAAATCCGTCTATAGATTAACCTGCCGAGGAGGTGTTGCCAA 1420
Db      1367  TAGCCAACTCGCGCGGCGGAAATCTCCATCATAGATAAATCTGCCGAGGAGGTGTTGCCAA 1426
Qy      1421  ATTATATATGCGCTTCAAGAGGAGCAGGCAAGCGAGCTTAAAGAACAAACCCCTTCAAGT 1480
Db      1427  ATTCTTACCGCTTCCCAAGGAGCAGCGCAAGCGCACTTAAGAACAAACCCCTTCAAGT 1486
Qy      1481  TCTTCGTTCCACCGT---CTCAGCAGTCTCCGAGGGGTGTTGGCTTAA 1524
Db      1487  TCTTCGTTCCACCTTTCGATCATCAGTCTATGAGGAGGTGGCTTAA 1533
```

```
RESULT 11
AF510854
LOCUS      3825 bp      DNA      linear      PLN 03-JUN-2002
DEFINITION Arachis hypogaea allergen Arah3/Arah4 gene, complete cds.
ACCESSION AF510854
VERSION    AF510854.1 GI:21314464
KEYWORDS
SOURCE     Arachis hypogaea (peanut)
ORGANISM   Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 3825)
AUTHORS   Viquez,O.M., Konan,K.N. and Dodo,H.W.
TITLE      Genomic characterization of the major peanut allergen genes, Arah3
and/or Arah4
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 3825)
AUTHORS   Viquez,O.M., Konan,K.N. and Dodo,H.W.
TITLE      Direct Submission
JOURNAL    Submitted (09-MAY-2002) Food and Animal Sciences, Alabama A&M
University, Meridian Street, 4900, Normal, AL 35762, USA
```

```
FEATURES             Location/Qualifiers
     source            1..3825
                     /organism="Arachis hypogaea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3818"
     TATA_signal       597..603
     mRNA              join(<661..949,1289..1581,1669..2313,2482..3825)
                     /product="allergen Arah3/Arah4"
     CDS                join(661..949,1289..1581,1669..2313,2482..2871)
                     /codon_start=1
                     /product="allergen Arah3/Arah4"
                     /protein_id="AAM46958.1"
                     /db_xref="GI:21314465"
                     /translations="MKLLALSVCFVLVLGASSIFRQOPBENACQFORLNAQRPDN
RISEGGYLETWPNNQEFCAGVALSLRLRNLRPFYSNAPQEIFIQGRGYFG
LIPFGCPSTVEEPAQGRHQSRPPRRFQGDQSQDQSHQKVRHFDGDLIAVPT
GVAFWMYNDHTDVAVSUTDNNNDQLDQPPRRFNLAGNHOEPLRYQOQSRRLS
PVSYPSTQPKQDEFEFSPRGQHRERAGOEENGNIFSGFTPEFLAQAQVDD
RQLQLRGENESEDGALVVRGRLILSPKRRQQRVERPDEEYDEYDEE
ERQDRRRGSGSGNGIEETICTASFKNIGRNSPDINYPQAGSLKATNELQNL
LILRWLGLSAEYLNALFVPHYNTNAHSIIYALRGRAHVQVDSNGDRVDEELQ
EGHLVVPQNFVAQGSQENFEYAFKTDSDRPSIANLAGENSFIDNLPEEVVANSYG
LPREAROLKNNNPFFVPPSPQSLRAVA"
                     <661..949
                     /number=1
                     1289..1581
     exon
     exon
```

```
exon      /number=2
           1669..2313
exon      /number=3
           2482..3825
           /number=4
           2872..3825
3'UTR
ORIGIN
Query Match      54.8%; Score 834.6; DB 8; Length 3825;
Best Local Similarity 77.1%; Pred. No. 2.6e-207;
Matches 1233; Conservative 0; Mismatches 69; Indels 297; Gaps 6;
Qy      214  CAGCAAGGAGGGGATACCTTTGGGTGTGATATTCCTCGTGTGTCTTAGACACTATGAAGAG 273
Db      1282  CATACAGGAGGGGATACCTTTGGTTTGTATTTCCCTGTGTCTTAGACACTATGAAGAG 1341
Qy      274  CTTCAACACAAGGTCGTGATCTCAGTCCCAAGACCAACAGAGCGTCTCAAGGAGAA 333
Db      1342  CTTGCAACAAGGAGCGCGCATCATGTCCTCAAGACCAACAGAGCGTTTCAAGGACAA 1401
Qy      334  GACCAAGGCAACAGCAGACAGATAGTCAACAGAGGTCACCGTTTCGATGAGGCGTAT 393
Db      1402  GACCAAGGCAACAGCAGACAGATAGTCAACAGAGGTCACCGTTTCGATGAGGCGTAT 1461
Qy      394  CTCATTGCAAGTCCCAACCGGTGTGTTCTCGCTCTACAACGACACGACACTGATGT 453
Db      1462  CTCATTGCAAGTCCCAACCGGTGTGTTCTCGATGTACACGACCATGACACTGATGT 1521
Qy      454  GTTGTGTTTCTTACTGACCAACACAGACAGTGTGACGTTGATCAGTTTCCCGAG- 512
Db      1522  GTTGTGTTTCTTACTGACCAACACAGACAGTGTGATCAGTTTCCCGAGG 1581
Qy      513  -----
Db      1582  GTATATATATGTTCAATCTCATTTTTCGATGTGCATATATATCATGAAGGA 1641
Qy      513  -----GAGATTCAATTTGGCTGGGAAACAGGAGCAAGAG 546
Db      1642  ATGGTGACAATTAATAATACATAACAGAGATTCAATTTGGCTGGGAAACAGGAGCAAGAG 1701
Qy      547  TTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGAGAGCTTACCATATAGCCCA 606
Db      1702  TTCTTAAGATACCAGCA-----ACAAAGCAGAGAGAGCTTACCATATAGCCCA 1752
Qy      607  TACAGCCCAAGTCAAGTACAGAGAGAGCGTGAATTTAGCCCTCGAGGAGCAGCAC 666
Db      1753  TACAGCCCGCAACTCAGCCTAAACAAGAGACCGTGAATTTAGCCCTCGAGGAGCAGCAC 1812
Qy      667  AGCCGAGAGAACGAGCAGGACAGAGAAAGAAACGAAAGGTGGAAACATCTTCAGCGGC 726
Db      1813  GGCCGAGAGAACGAGCAGGACAGAGAAACGAAAGGTGGAAACATCTTCAGCGGC 1872
Qy      727  TTCAGCCCGAGTCTCTGGAACAGCCTTCCAGGTTGACGACAGACAGATAGTGCAAAAC 786
Db      1873  TTCAGCCCGAGTCTCTGGAACAGCCTTCCAGGTTGACGACAGACAGATATTTGCAAAAC 1932
Qy      787  CTAAGAGGCGAGACCGCAGAGTGAAGAGAGGAGGCCATTTGTGCAGTGTAGGGAGGCGCTC 846
Db      1933  CTAAGAGGCGAGACCGCAGAGTGAAGAGAGGAGGCCATTTGTGCAGTGTAGGGAGGCGCTC 1992
Qy      847  AGAATCTTGAGCCCGAGATGAAGAGAG-----ACGTGCCGACGAAAGAGAG 891
Db      1993  AGAATCTTGAGCCCGAGATGAAGAGAGAGGCGAGCAGTATGAACGCTCCCGACGAAAGAGAG 2052
Qy      892  GAATACGATGAAGATGAATATGAATACGATGAAGAG-----CATAGAAGCGGT 939
Db      2053  GAATACGATGAAGATGAATATGAATATGAATGAAGAGAGAGGAGGCAACAAAGATGAAGCGGT 2112
Qy      940  GGCAGGGGAGCAGAGGCGAGGGGGAATGGTATTTGAAGAGACGATCTGCAACCGCAAGTGCT 999
Db      2113  GGCAGGGGAGCAGAGGCGAGGCGCAATGGCAATTTGAGGAGACCATCTGCAACCGCAAGTTTT 2172
Qy      1000  AAAAAACAATTTGGTAGAAACAGATCCCTCGATCTCAACCCCTCAAGCTGTTCACTC 1059
```



```
Db 2173 AAAAAGAACATTGGTAGAACAGATCCCTGACATCTACAACTCAAGCTGGTCACTC 2232
Qy 1060 AAAATCGCAACGATCTC-----AACCTTCTAATCTTAGTGGCTTGGACCTAGTGCT 1113
Db 2233 AAAATCGCAACGAGCTCCAGCTTAACCTTCTAATCTTAGTGGCTTGGACTTAGTGCT 2292
Qy 1114 GAATATCGAAATCTCTACAG-----1133
Db 2293 GAATATCGAAATCTCTACAGGTTTGTACTTTCTCTTTTCTATGTTTTTTTACGAATAT 2352
Qy 1134 -----1133
Db 2353 ATACTATCATTTTAAAGTATTAAATTTTAAATACTAATGCTATAAGAAATAAATGTTAAC 2412
Qy 1134 -----1133
Db 2413 TACTGTTGATTATATCGATAAAAGGATTACATTAGACATTTATTTTACATATTGAAATG 2472
Qy 1134 -----GAATGATTGTTTGTGCTCACTACAAACCAACGACACATCATATAT 1185
Db 2473 TGATTGAGAAATGATTTGTTGCTTCACTTACAAACGACGACACATCATATAT 2532
Qy 1186 CGATTGAGGGACGGGCTCACGTCGAAGTCGTGACAGCAACGCGCAACAGAGTGTACGAC 1245
Db 2533 GCATTGAGGGACGGGCTCACGTCGAAGTCGTGACAGCAACGCGCAACAGAGTGTACGAC 2592
Qy 1246 GAGAGGCTTCAAGAGGTCACGTCGTTGTTGGTGCACAGAACTTTCGCGCTCGCTGGAAG 1305
Db 2593 GAGAGGCTTCAAGAGGTCACGTCGTTGTTGGTGCACAGAACTTTCGCGCTCGCTGGAAG 2652
Qy 1306 TCCAGAGCAGAACTCCGTCATAGATAAAGCTGCGGAGAGGAGTGGTTGCAAAATTC 1425
Db 2653 TCCAGAGCAGAACTCCGTCATAGATAAAGCTGCGGAGAGGAGTGGTTGCAAAATTC 2772
Qy 1426 TATGGCTTCAAGAGGAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTCTTTC 1485
Db 2773 TATGGCTTCAAGAGGAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTCTTTC 2832
Qy 1486 GTTCCACCGCTCAGCAGTCTCGAGGGCTGTGCTTAA 1524
Db 2833 GTTCCACCGCTCAGCAGTCTCGAGGGCTGTGCTTAA 2871
```

RESULT 12

```
AY722687 1786 bp mRNA linear PLN 19-SEP-2004
LOCUS Arachis hypogaea arachin Ahy-3 mRNA, complete cds.
DEFINITION Arachis hypogaea arachin Ahy-3 mRNA, complete cds.
ACCESSION AY722687
VERSION AY722687.1 GI:52001222
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 1786)
Yan, Y., Wang, L. and Huang, S.
cDNA clone of peanut seed storage protein gene
Unpublished
2 (bases 1 to 1786)
Yan, Y., Wang, L. and Huang, S.
Direct Submission
Submitted (30-JUL-2004) Department of Biology and Biotechnology,
Sun Yat-sen University, Guangzhou, Guangdong 510275, China
JOURNAL Location/Qualifiers
FEATURES
source
/organism="Arachis hypogaea"
```

```

/mol_type="mRNA"
/db_xref="taxon:3818"
39..1493
/note="seed storage protein"
/codon_start=1
/product="arachin Ahy-3"
/protein_id="AAU21492.1"
/db_xref="GI:52001223"
/translation="MAKLLALSVCFCLVLGASSVTFRQGBENECQFORLNAQRDPN
CIESEGGYLETWNPNQEFQCAVALSRFLRNALRPFYSNAPQEIFYQSGGFYG
LIFFPGCTGFEEPIQSGEQFRPSRHFQDQSRFLDTHQKVHREGLDIAVPHGV
AFWIVNDQTDVAISVLHNSLNHLDQDPRFNLGAKQEQEFLRYQSRGSRQSGK
EQEOBOENEGNVSGFTEFLSHGVQVNEDEVRLNLSREREGEATVTVKGLSI
LVPEWRQSYQPGRGDKDFNNGIETICTATVKNNGIKSTADINPQAGSVRTVNE
LDLPILNRLGLSAEYGSIHEDAMFVPHYNNANSIYALHGGAHVQVDCNGRNVDE
EQEGSLVYPQNFVAASKSEHFDIVAFKTNRSASISNLAKGNSIMNLPEDVVAN
SYGLOYEQARQLKNNPNFTFLVPPQDSQMIRTV"
CDS
ORIGIN
Query Match 45.6%; Score 694.8; DB 8; Length 1786;
Best Local Similarity 70.6%; Pred. No. 1.1e-170;
Matches 1054; Conservative 0; Mismatches 197; Indels 141; Gaps 4;
Qy 2 GGCAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTTGACA 61
Db 109 GGCAGCAGGGGAGGAGATGAGTGCAGTTCAGCGCCTCAATGCCAGAGACCCGACA 168
Qy 62 ATGCAATTGAATCAGAGGCGGTTCATTTGAGACTTTGGAAACCCCAACAAACAGGAGTTGC 121
Db 169 ACTGCATTGAGTCTGAAGCGGTTCATTTGAGACTTTGGAAACCCCAACAAACAGGAGTTCC 228
Qy 122 AATGCGCGCGCTCGCGCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCTT 181
Db 229 AGTGCGCCGCGCTCGCGCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCTT 288
Qy 182 TCTACTTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGAGAGGGGATATCTTTGGGTGA 241
Db 289 TCTACTTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGAGAGGGGATATCTTTGGGTGA 348
Qy 242 TATTCCTCGTTGTCTTAGACACTATGAAGAGCTTCAACAGGCTTCAACAGGCTTCAAGT 301
Db 349 TATTCCTCGTTGTCTTAGACACTATGAAGAGCTTCAACAGGCTTCAAGT 402
Qy 302 CCCAAGACCAACCAAGACGCTTCCAGGAGAGAGACCAACCAAGCAGCAGGATAGTC 361
Db 403 TCCAAAGACCATCAAGACATTTTCAAGGCGAGGACCAACCAAGGCGCACTGGACATC 462
Qy 362 ACCAGAGGTGCACCGTTTCGATGAGGTTGATCTCATTCGAGTTCCACCGGTGTGCTT 421
Db 463 ACCAAGAGTTTACGGCTTCAGAGAGGTTGATCTCATCGTGTTCGCCACGGTGTGCTT 522
Qy 422 TCTGGCTTCAACCGACCAACGACACTGATGTTGTTGCTGTTTCTTCTTACTGACCAACA 481
Db 523 TCTGGATCTACAACGACCAAGACACTGATGTTGTTGCTGTTTCTGTTTACACCAACA 582
Qy 482 ACAACGACCAACGACTTGTATGATCTCCCGAGGATTCATTTGGCTGGGACACCGGAC 541
Db 583 GCTTCCACCAACCAACTTGCACAGTTTCCCGAGGAGTTCAATTTAGTGGAAAGCAAGAGC 642
Qy 542 AAGAGTTCTTAAGGTACACGACCAAGCAGACAAAGCAGACAAAGAGCTTACCATA 601
Db 643 AAGAATCTTACGATACCAAGCAACGAGTGTGTTCAAGGCCCAAAA-----689
Qy 602 GCCCATACAGCCCGCAAGTTCAGCTTAGACAAGAGGCGTGAATTTAGCCCTCGAGGAC 661
Db 690 -----689
Qy 662 AGCAGCGCGCAGAGAACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
Db 690 -----GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
Qy 722 GCGGCTTACGCGGAGTTCTCGGACAGAGCTTCCAGGTTGACGACAGACAGATAGTGC 781
```

Db	739	GCGGCTTCTCAACAGAAATTCCTCTCACATCGCTGCCAAGTGAACGA---GGACATAGTGA	795
QY	782	AAAACTAAGAGCGGAGACCGAGAGTGAAGAAGAGGAGCCATTGTCACAGTGAAGGGAG	841
Db	796	GGAACCTAAGAGGAGAAAACGAACGCGAGGAGCAAGAGGCAATTGTCACAGTAAAGGAG	855
QY	842	GCCTCAGAAATCTTGAGCCCAAGATAGAAAGAGAGAGTGCAGAGAGAGAGGAATACGATG	901
Db	856	GACTTAGCATCTTAGTTTCCACAGAGTGGAGGAGAGCTACCAACAA-----	902
QY	902	AAGATGAATATGATACGATGAAGAGATAGAAGCGCTGCGAGGGAAGCAGAGGCGAGG	961
Db	903	-----	927
QY	962	GGAATGATATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAACATTTGGTAGAAACA	1021
Db	928	ACAAAGCATTTGAAGAAACCATTTGCACTGCACTGTTAAGATGAACATTTGGCAATCCA	987
QY	1022	GATCCCTGACATCTCAACCCCTCAAGCTGGTTTCACTCAAAAAGTGCACACGATCTCAACC	1081
Db	988	CATCTGCTGATATCTCAACCCCTCAAGCTGGTAGCGTCAGAACTGTAAACGAACTTGACC	1047
QY	1082	TTCTAATACCTTAGTGGCTTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCAT	1141
Db	1048	TCCCAATCTCAATCGACTCGGACTCAGTGCAGATGGATCCATTTCACTCGGGATGCAA	1107
QY	1142	TGTTTGTGCTCACTCAACCAACCAACGACACAGCATCATATCGATTGAGGGGACGGG	1201
Db	1108	TGTTTGTCTCACTACACATGAAACGCAACAGCATGATATGATTCACGCGAGGAG	1167
QY	1202	CTCAGTGCAGTGTGGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGG	1261
Db	1168	CTCATGTCCAAAGTGGTGGAGTGCATGGCAATAGAGTGTTCGACGAGGAGCTTCAAGAGG	1227
QY	1262	GTCAAGTGTGTGGTGGCCACAACTTCGCGCTCGCTGGAAGTCCGACGAGGCAACT	1321
Db	1228	GTCAATCGCTGGTGGTGGCCACAAACTTCGCGCTGCTGGAAGTCCGACGAGGAGCACT	1287
QY	1322	TGCAATACGTGGCAATTCAGACAGACTCAAGGCCACAGCATAGCAACCTCGCGGGTGA	1381
Db	1288	TCTTGTAGTGGCAATTCAGACAACTCAAGGCCAGCATATCCAACCTTGTGGCAAA	1347
QY	1382	ACTCGCTATAGATAACCTGCGGAGAGGTGGTTGCAATTCATATGCGCTTCCAAAGG	1441
Db	1348	ATTCTACATGTGGAATCTGCCGGAAGTGTGGTTGCAATTCATATGCGCTTACAAATAG	1407
QY	1442	AGCAGGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACC	1493
Db	1408	AGCAAGCAAGGCAACTCAAGAACAAATAACCCCTTCACGTTCTTGGTTCCACC	1459
RESULT 13			
AF487543	Arachis hypogaea trypsin inhibitor mRNA linear PLN 02-APR-2004		
LOCUS	Arachis hypogaea trypsin inhibitor mRNA, partial cds.		
DEFINITION	Arachis hypogaea trypsin inhibitor mRNA, partial cds.		
ACCESSION	AF487543		
VERSION	AF487543.1 GI:22135347		
KEYWORDS	Arachis hypogaea (peanut)		
SOURCE	Arachis hypogaea		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
REFERENCE	1 (bases 1 to 657)		
AUTHORS	Dodo,H.W., Viquez,O.M., Maleki,S.J. and Konan,K.N.		
TITLE	cDNA clone of a putative peanut (Arachis hypogaea L.) trypsin inhibitor has homology with peanut allergens Ara h 3 and Ara h 4		
JOURNAL	J. Agric. Food Chem. 52 (5), 1404-1409 (2004)		
PUBMED	14995153		
REFERENCE	2 (bases 1 to 657)		
AUTHORS	Dodo,H.W. and Viquez,O.M.		
TITLE	Direct Submission		

JOURNAL	Submitted (25-FEB-2002) Food and Animal Sciences, Alabama A&M University, Meridian Street, 4900, Normal, AL 35762, USA				
FEATURES	Location/Qualifiers				
source	1..657				
	/organism="Arachis hypogaea"				
	/mol_type="mRNA"				
	/db_xref="taxon:3818"				
	<1..>657				
	/codon_start=1				
	/product="trypsin inhibitor"				
	/protein_id="AAM93157.1"				
	/db_xref="GI:22135348"				
	/translation="NYLHMLLALSVCFLVLGASSISFRQPEENACQFQRLNARQRP DNRISEGYITWPNNEQECAGVALSLVLRNALRRPFYSNAPQEIFIQQRGY FGLIPGCPSTVEEPAQQRHQSOPRRFQGSQSOQSDSHQKVRPDRGDLIAV PTGVAFMYNDHDTDVVASLTDITNNNDNQDQPPRRFLNAGNHEQFLRYQQQSRRR S"				
CDS					
	Query Match	34.1%;	Score 519.2;	DB 8;	Length 657;
	Best Local Similarity	93.4%;	Pred. No. 1e-124;		
	Matches 542;	Conservative 0;	Mismatches 38;	Indels 0;	Gaps 0;
QY	2	GGCAGAACCGAGAGAGACGGTCCAGTTCACGCGCTCAATGCGAGACCTGACA	61		
Db	77	GGCAGCAGCGGAGGAAATGCGTCCAGTTCACGCGCTCAATGCGCAAGGCGCTGACA	136		
QY	62	ATCGCATTAATCAGAGGCGGTACATTGAGACTTTGGAACCCCAACACAGGAGTTGG	121		
Db	137	ACCGCATTAATCGGAGGCGGTACATTGAGACTTTGGAACCCCAACACAGGAGTTGG	196		
QY	122	AATGCGCGGGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAAGGCCCTTCGTAGGCTT	181		
Db	197	AATGCGCGGGCTCGCCCTCTCGCGCTTAGTCTCTCGCGCAAGGCCCTTCGTAGGCTT	256		
QY	182	TCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAATCTTTGGGTTGA	241		
Db	257	TCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAATCTTTGGTTTGA	316		
QY	242	TATTCCTCGTTGCTTAGACACTATGAAGAGCCTCACACACAAGTGTGCTGATCTCAGT	301		
Db	317	TATTCCTCGTTGCTTAGACACTATGAAGAGCCTGCAACAAGAGCGCCGACATCAGT	376		
QY	302	CCCAAGACCAACCAAGAGCGTCTCCAAGGAGAGACCAAGCCCAACAGCAGATAGTC	361		
Db	377	CCCAAGACCAACCAAGAGCGTCTCCAAGGAGAGACCAAGCCCAACAGCAGATAGTC	436		
QY	362	ACCAGAGGTGACCGTTTCGATGAGGGTGATCTCATTCGAGTTCACCGGTGTTGGTTT	421		
Db	437	ACCAGAGGTGACCGTTTCGATGAGGGTGATCTCATTCGAGTTCACCGGTGTTGGTTT	496		
QY	422	TCGTGCTCTACAAACGACGACATGATGTTGTTGCTTCTTCTTACTGACACCAACA	481		
Db	497	TCGTGATGTACAAACGACGACATGACATGATGTTGTTGCTTCTTCTTACTGACACCAACA	556		
QY	482	ACAACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGAGC	541		
Db	557	ACAACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGAGC	616		
QY	542	AAGAGTCTTAAGGTACAGCAACAAAGCAGCAAAAGCAG	581		
Db	617	AAGAGTCTTAAGATACCAGCAACAAAGCAGCAAAAG	656		
RESULT 14					
BD175913	BD175913				
LOCUS	1446 bp DNA linear PAT 18-MAR-2003				
DEFINITION	Regulation of soybean seed protein gene of definite class.				
ACCESSION	BD175913				
VERSION	BD175913.1 GI:29121615				
KEYWORDS	JP 2002262894-A/13.				
SOURCE	unidentified				
ORGANISM	unidentified				

REFERENCE	unclassified.	
AUTHORS	1. (bases 1 to 1446)	
TITLE	Kinney,A.J. and Fader,G.M.	
JOURNAL	Regulation of soybean seed protein gene of definite class	
COMMENT	Patent: JP 2002262894-A 13 17-SEP-2002; EI DU PONT DE NEMOURS AND CO OS Unidentified PN JP 2002262894-A/13 PD 17-SEP-2002 PF 28-FEB-2002 JP 2002053791 PR 14-JUN-1996 US 60/019940 PI ANTHONY J KINNEY,GARY MICHAEL FADER PC C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC Strandedness: Single; CC Topology: linear; CC Regulation of soybean seed protein gene of definite class FH Key Location/Qualifiers FT source 1. .1446 FT Location/Qualifiers FT /organism='Unidentified'. FT 1. .1446 FT /organism='unidentified' FT /mol_type='genomic DNA' FT /db_xref='taxon:32644'	
FEATURES		
source		
ORIGIN		
Query Match	34.0%; Score 517.4; DB 6; Length 1446;	
Best Local Similarity	62.8%; Pred. No. 3.3e-124;	
Matches	954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;	
QY	5 AGCAACCGGAGGAAACGCGTCCAGTTCAGCGCCCTCAATGCGCAGAGACCTGACAATC	64
DB		
QY	71 AGCAGCCAGCAAAAGAGTGGCCAGATCCACGCCCTCAATGCCCTAAACCGGATAACC	130
DB		
QY	65 GCATTGAATCAGAGGGCGGTTACATTGAGAGCTTGGAAACCCCAACCAACAGGAGTTCCGAAT	124
DB		
QY	131 GTATAGAGTCAGAAGTGGCTTCATTGAGACATGGNACCCTTAACAACAGCCATTCCAGT	190
DB		
QY	125 GCGCGGGTGGCCCTCTCTCGCTTAGTCTCGCCCGCAAGCCCTTCGTAGGCTTTCT	184
DB		
QY	191 GTGCGGTGTGGCCCTCTCTCGCTGACCCCTCAACCGCAAGCCCTTCGAGACCTTCCT	250
DB		
QY	185 ACTCCAATGTCCCGAGGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTGATAT	244
DB		
QY	251 ACACCAAGCTCCCGAGGATCTACATCAACAAAGTAGTGGTATTTTGGCATGATAT	310
DB		
QY	245 TCCCTGTTGTCTTAGACACTATGAAGAGCTTCACACAAAGTCTCGATCTCAGTCCC	304
DB		
QY	311 TCCGGGTGTCTTAGCACATTGAAGAGCT-----	342
DB		
QY	305 AAGACCACCAAGAGCTCTCAAGGAGAAAGACCAAGCCCAACAGCAACGAGATAGTCAAC	364
DB		
QY	343 -----CAACAAAGAGCAAAAGCAGCGGCCCAAGAGCCGTCAAC	382
DB		
QY	365 AAGAGTGACCGTTTCGATGAGGGTATCTCATGTCAGTTCACCGGTGTTGCTTTCT	424
DB		
QY	383 AGAAGATCTATCACTTCAGAGAGGGTGAITTGATGTCAGTGCACACCGGTTTGGCATACT	442
DB		
QY	425 GGCTCTCAACAGCAGCACACTGATGTGTGTGTTTCTCTTACTGACACCAACAACA	484
DB		
QY	443 GGATGTACAACTGAAGACACTCTCTGTGTGTCGGCTTCTCTTATTGACACCAACAGCT	502
DB		
QY	485 ACACAACAGCTTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGAAACCGAGCAAG	544
DB		
QY	503 TCCAGAACAGCTCGACAGATGCTAGGAGATTCTATCTTGTCTGGGAACCAAGAGCAAG	562
DB		
QY	545 AGTTCTTAAGTACAGCAACAAAGCAGACAAAGCAGACGAGCTTACCATATAGCC	604
DB		
QY	563 AGTTCTTACAGTATCAGCCACAGAGCAGCAAG-----	595
DB		
QY	605 CATACAGCCCGCAAGTGCAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGACAGC	664
DB		
QY	596 -----GAGGTACTC	604

QY	665 ACAGCCGACAGAAACGAGCAGGACAAGAGAAAGAAACGAGGTGAAACATCTTCAGCG	724
DB		
QY	605 AAAGCCAGAAAGGAAAGCGTCAGCAAGAAAGAAAGAAAGGAGCGAGCATATTGAGTG	664
DB		
QY	725 GCTTTCAGCGCGGAGTTCTCTGGAACAAGCCCTTCCAGGTTGACGACAGACATAGTCAAA	784
DB		
QY	665 GCTTCGCCCCGGAATTTCTTGGAAACATGCGTTC---GTCGTGGACAGGCAGATAGTGA	721
DB		
QY	785 ACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGGCCATTGTGACAGTGGGGAGGCC	844
DB		
QY	722 AGCTCAAGGTGAGAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	781
DB		
QY	845 TCAGATCTTGAGCCCGCAGATAGAAAGAGAGCGTCCGACGAGAGAGGAGGAGGAGGAG	904
DB		
QY	782 TCAGCGTGAATAAGCCCAACCCAGCAAGAGAGCAGCAACAAGAGCCGAGGAGAGGAG	841
DB		
QY	905 ATCAATATGAATACGATGAAGAGGATAGAAGGCGTGTGCGAGGGAAGCAGAGCGAGG	964
DB		
QY	842 CAGATTGTGACGAGAAAGACAAACAT-----TCCCAAGGCCAAAGCAGAA	886
DB		
QY	965 ATGCTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAAGACATTGGTAGAAACAG	1024
DB		
QY	887 ATGCGATTGACGAGACCAATTTGCAATGAGACTTCCGCCACAACATTGGCCAGACTTCA	946
DB		
QY	1025 CCCTGACATCTACAAACCTCAAGCTGTGTTCACTCAAAAAGTCCCAAGGATCTCAACCTTC	1084
DB		
QY	947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCCGTACCGACTCTCGACTTC	1006
DB		
QY	1085 TAATACCTTAGGTGCTGGACCTAGTGTGAATATGMAATCTCTACAGGAATGCATTGT	1144
DB		
QY	1007 CAGCCCTCTGCTGGCTCAAACTCAGTGCACAGTTCGATCCTCCGCAAGATGCTATGT	1066
DB		
QY	1145 TTGTCCTCACTACACCAACGACACAGCATCATATATCGATTTGAGGGGACGGGCTC	1204
DB		
QY	1067 TCGTGCCACACTACAACCTGAAACGCAACAGCATATATACGCAATTGAAATGGAGCG	1126
DB		
QY	1205 AGTGCAAGTCTGGAGCAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC	1264
DB		
QY	1127 TGTACAAAGTGGTGAATTCGAATGGTGGAGAGTGTGTTGATGGAGAGCTGCAAGAGGG	1186
DB		
QY	1265 ACCTGCTGTGGTGGCCACAGAACTTCGCGCTCGCTGGAAGTCCAGAGCGAGAACTTCG	1324
DB		
QY	1187 AGTGTTAATTTGTCACAAAACTTTGCGTGGCTGCAAGTCAAGAGCGACACTTCG	1246
DB		
QY	1325 AATACTGGCATTTCAAGACAGACTCAAGGCCAGCAGTAGCCAACTCGCGGTGAATACT	1384
DB		
QY	1247 AGTATGTTTCAATTTCAAGACCAATGATAGACCCCTCGATCGCAACCTTGCAGGTGCA	1306
DB		
QY	1385 CGTCAATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCTTCCAAAGGGAG	1444
DB		
QY	1307 CATTTGTTGAACGCAATTTGCGGAGGAAAGTATTTCAGCAAACTTTTAACTTAAGGAG	1366
DB		
QY	1445 AGCAAGGAGCTTTAAGAAACAACCCCTTCAAGTTCTTTCGTTCCACCGTCTCAGCAGT	1504
DB		
QY	1367 AGCCAGGAGGTCAAGAAACAACCCCTTTCAGTTCTTTCGTTTCCACCTTCAAGAGTCT	1426
DB		
QY	1505 CTCGAGGGCTGTGGCTTA 1523	
DB		
QY	1427 AGAGGAGAGTTGTGGCTTA 1445	
DB		
RESULT 15		
LOCUS	AR202578	
DEFINITION	Sequence 13 from patent US 6362399.	
ACCESSION	AR202578	
VERSION	AR202578.1	
KEYWORDS	GI:20257117	
SOURCE	Unknown.	
ORGANISM	Unknow.	
REFERENCE	1 (bases 1 to 1446)	

AR202578	1446 bp	DNA	linear	PAT 20-APR-2002
Sequence 13 from patent US 6362399.				
AR202578				
AR202578.1				
GI:20257117				
Unknown.				
Unknow.				
Unclassified.				
1 (bases 1 to 1446)				

AUTHORS		Kinney, A. John. and Fader, G. Michael.	
TITLE		Suppression of specific classes of soybean seed protein genes	
JOURNAL		Patent: US 6362399-A 13 26-MAR-2002;	
FEATURES	source	Location/Qualifiers	
	1. .1446	/organism="unknown"	
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match		34.0%;	Score 517.4; DB 6; Length 1446;
Best Local Similarity		62.8%;	Pred. No. 3.3e-124;
Matches		954; Conservative	0; Mismatches 421; Indels 144; Gaps 4;
QY	5	AGCAACGGAGGAGAACGCGTGCAGTTCACAGCGCCCTCAATGCGCAGAGACCTTGACAATC	64
DB	71	AGCAGCCACAGCAACAGAGTGCAGATCCAAAGCCCTCAATGCCCTAAACCCGATACC	130
QY	65	GCATTGAATCAGAGGGCGGTTCATTGAGACTTTGGAAACCCCAACACAGGAGTTGCAAT	124
DB	131	GTATAGAGTCAGAAAGGTGGCTTCATTGAGACATGGAACCCCTAAACAAGCCATTCCAGT	190
QY	125	GCGCCGGTGCCTCTCTCGCTTAGTCTCCGCCGACGCGCTTCGTAGGCTTTCT	184
DB	191	GTGCCGGTGTGGCCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTCGCGAGACCTTCCT	250
QY	185	ACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTTCATAT	244
DB	251	ACACCAACGCTCCCCAGGAGATCTACATCCCAACAGGTAGTGTATTTTGGCATGATAT	310
QY	245	TCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAGTCCC	304
DB	311	TCGCCGGTGTCTCTAGCACATTTGAAGGCCT-----	342
QY	305	AAAGACCACCAAGACGTCTCCAGGAGAGAACCAAGCCCAACAGCAGATAGTCACC	364
DB	343	-----CAACAAAGAGGACAAAGCAGCAGGCCCCCAAGACCGTCACC	382
QY	365	AGAAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCACCACGGTGTGCTTTCT	424
DB	383	AGAAAGATCTATCACTTCAGAGAGGGTGTATTTGATTCAGTGCACACGGTTTTCATCT	442
QY	425	GGCTCTCAACACGACACGACACTGATGTTGTTGCTTTCTTTACTTGACACCAACAAC	484
DB	443	GGATGTACAAATGAAGACACTCCTGTTGTTGCGGTTTCTCTTATTGACACCAACAGCT	502
QY	485	ACACAAACCACTGTATCAGTTCCTCCAGGAGATTCATTTGGCTGGGACACACCGGCAAG	544
DB	503	TCAGAAACCACTGCGACCATGCTAGGAGATTTCTATCTTGCTGGGAACCAAGAGCAAG	562
QY	545	AGTTCTTAAGGTACCAAGCAACAAAGCAGACAGCAAGAGAGCTTACCATATAGCC	604
DB	563	AGTTCTTACAGTATCAGCCACAGACGACGACG-----	595
QY	605	CATACAGCCCGCAAGTGCAGCTAGACAGAGAGCGTGAATTTAGCCCTTCGAGGACAGC	664
DB	596	-----GAGGTACTC	604
QY	665	ACAGCCGACAGACGAGCAGGACAGAAAGAAACGAGGTGGAAACATCTTCAGCG	724
DB	605	AAAGCCAGAAAGGAAAGCGTTCAGCAAGAAAGAAACGAAAGGAGGAGCATATTGAGTG	664
QY	725	GCTTCAGCGCGGAGTTCCTGGAAACAGCCCTTCAGGTTTGACACAGACAGATAGTGCAA	784
DB	665	GCTTCGCCCCGGAATCTTGGACATGCGTTC---GTGCTGGACAGGCAGATAGTGAGAA	721
QY	785	ACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGGGGAGGCC	844
DB	722	AGCTCAAGAGGTGAGAACGAAGAGGAGAGAGAGGGTGCATTGTGCACAGTGAAGAGGTC	781
QY	845	TCAGAACTTTGAGCCAGATAGAAAGAGACGTGCCGACGAGAGAGGAGATACGATGAG	904
DB	782	TCAGCGTGAATAGCCCAACCCAGGAGAGCAGCAACAAAGCCCGAGGAGAGGAGAGC	841

Search completed: August 24, 2005, 00:57:21
Job time : 4522.08 secs

QY	905	ATGAATATGAATACGATGAAGAGGATAGAAAGCGGTGGCAGGGGAAAGCAGAGCGCAGGGGA	964
DB	842	CAGATTGTGACGAGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA	886
QY	965	ATGGTATTGAAGAGAGGATCTGCACCGCAAGTGTCTAAAGAACATTTGTAGNAACAGAT	1024
DB	887	ATGGCATTTGACGAGACCAATTTGCACAAATGAGACTTTGCCACCAACATTTGCCAGACTTCAT	946
QY	1025	CCCTTGACATCTACAAACCTCAAGCTGGTTCACCTCAAAAACCTGCCAACGATCTCAACCTTC	1084
DB	947	CACCTGACATCTCAACCTCAAGCTGGTAGCATCAACCGTACCAGCTCGACTTC	1006
QY	1085	TAATCTTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGAAATGCAATGT	1144
DB	1007	CAGCCCTCTCGTGGCTCAAACTCAGTGGCCAGTTTGGATCACTCCGCAAGAATGCTATGT	1066
QY	1145	TTGTGCTCACTACAAACACCAACGACACAGCATCATATATCGATTGAGGGGACCGGCTC	1204
DB	1067	TCGTGCCACACTACAACTGAACCGCAACAGCATATATACGATTTGAATGGACGGGCAT	1126
QY	1205	AGTGTCAAGTCTGTGACAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC	1264
DB	1127	TGTTACAGTGGTGAATTGCAATGGTGAGAGAGTGTGATGGAGAGCTGCAAGAGGGAC	1186
QY	1265	AGTGTCTTGTGTGCCACAGAACTTGGCCGTGCTGGAAAGTCCCAGAGCGAGAACTTCG	1324
DB	1187	AGTGTGTTAATTGTGCCACAAACTTTTGGGTGGCTGCAAGATCACAGAGCGCAACTTCG	1246
QY	1325	AATACGTGCAATTCAGACAGACTCAGGCCCAGACATAGCCAACTCGCCGGTGAACACT	1384
DB	1247	AGTATGTTTTCATTCAAGACCAATGATAGACCTCGATCGGCAACCTTCAGGGTGCAACT	1306
QY	1385	CGTCTATAGATAACCTGCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGGGAGC	1444
DB	1307	CATTGTTGAACGATTTGCCGAGGAGAGTGAATTCAGCAAACTTTTAACCTTAAGGAGCAGC	1366
QY	1445	AGGCAAGGAGCTTTAAGAAACAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGT	1504
DB	1367	AGGCCAGGCGAGTCAAGAACAAACCCCTTCAGCTTCCTGTTCCACCTAAGGAGTCTC	1426
QY	1505	CTCCGAGGCGTGTGGCTTA	1523
DB	1427	AGAGGAGAGTTGTGGCTTA	1445

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 570.653 Seconds
(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 CGGCGAACCGAGGAGGAGAA.....ctccgaggcgctgtggtcttaa 1524

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1524	100.0	1524	2	Az06384 Peanut al
2	1524	100.0	1524	2	Az22280 Ara h 3 a
3	1524	100.0	1524	8	Abx70612 Peanut Ar
4	1524	100.0	1524	10	Adg27543 Peanut CD
5	1520.8	99.8	1524	4	Aas08540 DNA encod
6	1391.2	91.3	1855	4	Aaf90338 Peanut al
7	517.4	34.0	1446	2	Aav17567 Coding se
8	517.4	34.0	1446	6	Abss5193 Glycine m
9	517.4	34.0	1446	10	Adh89252 G. max gl
10	517.4	34.0	1446	12	Adg43987 G. max gl
11	513	33.7	1488	2	Aav17565 Coding se
12	513	33.7	1488	6	Abss5191 Glycine m
13	513	33.7	1488	10	Adh89244 G. max gl
14	513	33.7	1488	12	Adg43979 G. max gly
15	513	33.7	1743	3	Aaz92638 cDNA enco
16	505.6	33.2	1458	6	Abss5192 Glycine m
17	505	33.1	1746	4	Aad17536 Soybean g
18	504	33.1	1458	2	Aav17566 Coding se
19	504	33.1	1458	10	Adh89246 G. max gl
20	504	33.1	1458	12	Adg43981

21	502	32.9	1712	2	AAQ05359	Glycinin
22	498.6	32.7	1746	2	AAQ05360	Glycinin
23	224	14.7	1554	6	ABSS5194	Glycine m
24	198	13.0	1551	2	AAV17569	Coding se
25	198	13.0	1551	10	ADH89250	G. max gl
26	198	13.0	1551	12	ADG43985	G. max gl
27	198	13.0	1786	2	AAQ05357	Glycinin
28	193.2	12.7	3527	11	ADL90167	DNA encod
29	189	12.4	1374	12	ADO43060	Cashew nu
30	188.8	12.4	1786	1	AAAN60939	Sequence
31	172.2	11.3	1689	10	ADH89248	G. max gl
32	172.2	11.3	1689	12	ADG43983	G. max gl
33	170.6	11.2	1899	2	AAQ05358	Glycinin
34	165.2	10.8	1646	1	AAH82246	Rice stor
35	162	10.6	1500	10	ADC08274	Rice DNA
36	153.2	10.1	1689	2	AAV17568	Coding se
37	153.2	10.1	1689	6	ABSS5195	Glycine m
38	151.8	10.0	1482	10	ADH89254	Sunflower
39	151.8	10.0	1482	12	ADG43989	H. annuus
40	149.8	9.8	1896	1	AAH60940	Sequence
41	147.8	9.7	1512	10	ADC07963	Rice DNA
42	144.2	9.5	2058	12	ADJ44956	Plant cDN
43	143	9.4	1512	10	ADC07957	Rice DNA
44	142	9.3	1706	2	AAH37335	Coffee et
45	140.2	9.2	1685	2	AAQ54819	Oat globi

ALIGNMENTS

RESULT 1
AAZ06384
ID AAZ06384 standard; DNA; 1524 BP.
XX
AC AAZ06384;
XX
DT 17-OCT-2003 (revised)
DT 09-NOV-1999 (first entry)
XX
DE Peanut allergen, Ara h 3.
XX
KW allergy; immune response; transgenic; allergen; epitope;
KW immunoglobulin E; Ig E; binding site; peanut; ds.
XX
OS Arachis hypogaea.
XX
FH Key Location/Qualifiers
FT CDS 1..1524
FT /tag= a
FT /product= "Ara h 3"
FT /note= "First three amino acids not coded for"
XX
XX WO9938978-A1.
XX
XX 05-AUG-1999.
XX
XX 29-JAN-1999; 99WO-US002031.
XX
XX 31-JAN-1998; 98US-0073283P.
XX 13-FEB-1998; 98US-0074590P.
XX 13-FEB-1998; 98US-0074624P.
XX 13-FEB-1998; 98US-0074633P.
XX 27-AUG-1998; 98US-00141220.
XX
XX (UYAR-) UNIV ARKANSAS.
XX (UYNV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX (SOSI/) SOSIN H.
XX
XX Sosin H, Bannon GA, Burks AW, Sampson HA;
XX WPI; 1999-479189/40.
XX P-PSDB; AAV15246.

PT	Modified allergen with reduced IgE binding, useful for treating e.g. allergies.	
PT		
XX	Dislosure; Page 38-39; 46pp; English.	
PS		
XX	This is the nucleotide sequence of the Ara h 3 protein from Arachis hypogaea. The Ara h 3 protein has 4 IgE (Immunoglobulin E) binding epitopes, one of which is immunodominant (AAV15281). By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response. (Updated on 17-OCT-2003 to standardise OS field)	
XX		
SQ	Sequence 1524 BP; 455 A; 390 C; 398 G; 281 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1524; DB 2; Length 1524;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CGGCAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60	781 CAAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGGGAGCCATTGTGTGACGTAGGGGA 840
Db	1 CGGCAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60	841 GGCCTCAGAAATCTTGAGCCACAGATGAAAGAGACGTGCCGACGAAGAGGAATACGAT 900
Qy	61 AATCGCATTTGAATCAGAGGGCGGTACATTTGAGACTTGGAAACCCCAACACGAGGATTC 120	841 GGCCTCAGAAATCTTGAGCCACAGATGAAAGAGACGTGCCGACGAAGAGGAATACGAT 900
Db	61 AATCGCATTTGAATCAGAGGGCGGTACATTTGAGACTTGGAAACCCCAACACGAGGATTC 120	901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAGAGAGGCGAG 960
Qy	121 GAATGCGCGCGGTGCGCCCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180	901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAGAGAGGCGAG 960
Db	121 GAATGCGCGCGGTGCGCCCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180	961 GGAATGATTTGAAGAGACGATCTTGACCCGCAAGTGTCTAAAGAGAACTTGGTAGAATC 1020
Qy	181 TTCTACTCCAATGTCCTCCAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGGTTG 240	961 GGAATGATTTGAAGAGACGATCTTGACCCGCAAGTGTCTAAAGAGAACTTGGTAGAATC 1020
Db	181 TTCTACTCCAATGTCCTCCAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGGTTG 240	1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTTCACTAAAACCTGCCAACGATCTCAAC 1080
Qy	241 ATATTCCTCGTTGTCCTAGACACTATGAAGAGCCTCACACAAGGTGTCGATCTCAG 300	1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTTCACTAAAACCTGCCAACGATCTCAAC 1080
Db	241 ATATTCCTCGTTGTCCTAGACACTATGAAGAGCCTCACACAAGGTGTCGATCTCAG 300	1081 CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Qy	301 TCCCAAGACACCAAGACCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360	1081 CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Db	301 TCCCAAGACACCAAGACCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360	1141 TTGTTTGTGCTCATCAACAACCAACGACACAGCATCATATATCGATTGAGGGACCG 1200
Qy	361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCACTTCCACCGGTGTGCT 420	1141 TTGTTTGTGCTCATCAACAACCAACGACACAGCATCATATATCGATTGAGGGACCG 1200
Db	361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCACTTCCACCGGTGTGCT 420	1201 GCTCAGTCCAAAGTCTGTGACAGCAACGCAACAGAGTGTACGAGGAGCTTCAAGAG 1260
Qy	421 TTCTGGCTCTAACACGACGACGACTGATGTTGTTGCTGTTTCTTCTTACTGACCAAC 480	1201 GCTCAGTCCAAAGTCTGTGACAGCAACGCAACAGAGTGTACGAGGAGCTTCAAGAG 1260
Db	421 TTCTGGCTCTAACACGACGACGACTGATGTTGTTGCTGTTTCTTCTTACTGACCAAC 480	1261 GGTCAAGTCTTGTGTGTGCCACAGAACTTCGCGCTGCTGGAAAGTCCCGAGCGAGAAC 1320
Qy	481 AACACGACAAACAGCTTGATCAGTTCCCGAGGAGATTCATTTGGTGGGAACCGGAG 540	1261 GGTCAAGTCTTGTGTGTGCCACAGAACTTCGCGCTGCTGGAAAGTCCCGAGCGAGAAC 1320
Db	481 AACACGACAAACAGCTTGATCAGTTCCCGAGGAGATTCATTTGGTGGGAACCGGAG 540	1321 TTGCAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCACCTCCCGGTGAA 1380
Qy	541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600	1321 TTGCAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCACCTCCCGGTGAA 1380
Db	541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600	1381 AACTCGTCTATAGATAAACCCTGCGGAGGAGTGGTTGCAAAATTCATATGGCTCCAAAGG 1440
Qy	601 AGCCCATACAGCCCGCAAGTACGCTAGCAAGAGAGCGGTGAATTTAGCCCTCGAGGA 660	1381 AACTCGTCTATAGATAAACCCTGCGGAGGAGTGGTTGCAAAATTCATATGGCTCCAAAGG 1440
Db	601 AGCCCATACAGCCCGCAAGTACGCTAGCAAGAGAGCGGTGAATTTAGCCCTCGAGGA 660	1441 GAGCAGGCAAGCGACTTAAAGAACAAACACCCTTCAAGTCTTCCACCGTCTCAG 1500
Qy	661 CAGCAGACCGCGCAGAGAACGAGCAGACGAAGAGAAACGAAGTGGAAACATCTTTC 720	1441 GAGCAGGCAAGCGACTTAAAGAACAAACACCCTTCAAGTCTTCCACCGTCTCAG 1500
Db	661 CAGCAGACCGCGCAGAGAACGAGCAGACGAAGAGAAACGAAGTGGAAACATCTTTC 720	1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
Qy	721 AGCGGCTTACCGCGGAGTTCTTGGAAACAGCCCTCCAGGTTCCAGCAGACAGATAGTG 780	1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
Db	721 AGCGGCTTACCGCGGAGTTCTTGGAAACAGCCCTCCAGGTTCCAGCAGACAGATAGTG 780	
Qy	781 CAAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGGAGGCCATTGTGACAGTGGGGGA 840	

RESULT 2	
AAZ22280	
ID	AAZ22280 standard; DNA; 1524 BP.
XX	AAZ22280;
XX	AC
XX	17-OCT-2003 (revised)
DT	06-DEC-1999 (first entry)
DT	
XX	Ara h 3 allergen encoding DNA.
DE	
XX	Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW	allergic reaction; ss.
XX	Arachis hypogaea.
OS	
XX	
FH	Key Location/Qualifiers
FT	1..1524
FT	/tag= a
FT	/transl_except= (pos:490..491, aa:Glu, Phe)
FT	/note= "there is an apparent 6 nucleotide deletion which encodes for residues Glu and Phe respectively"

FT	/transl_except= (pos:1105..1107, aa:Leu)
FT	/product= "Ara H 3 allergen"
FT	/note= "does not encode the first three residues in the corresponding protein sequence"
WT	WO9945961-A1.
PN	
XX	
PD	16-SEP-1999.
XX	
PP	12-MAR-1999; 99WO-US0005494.
XX	
PR	12-MAR-1998; 98US-0077763P.
PR	11-MAR-1999; 99US-00077763.
XX	
PA	(UYAR-) UNIV ARKANSAS.
XX	
PI	Burks W, Helm RM, Cockrell G, Bannion CA, Stanley JS, Shin DS;
PI	Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX	
DR	WPI; 1999-551218/46.
P-	PSDB; AAY40912.
XX	
PT	Tertiary structure of peanut allergen Ara h 1 for protection of a host animal from allergic reaction.
XX	
PS	Disclosure; Page 185; 193pp; English.
CC	The invention provides a tertiary structure for the peanut allergen Ara H 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding epitopes. The invention also provides an isolated recombinant peanut allergen designated Ara h 3 and a nucleotide molecule encoding the peanut allergen Ara h 3. Molecules of the invention are used to protect a host animal from allergic reaction, particularly using a modified allergen which is less reactive with IgE. The invention may also be used to ensure that the allergen is not introduced into genetically modified food. The present sequence represents a DNA encoding the Ara h 3 allergen. (Updated on 17-OCT-2003 to standardise OS field)
XX	
SQ	Sequence 1524 BP; 455 A; 390 C; 398 G; 281 T; 0 U; 0 Other;
Query Match	100.0%; Score 1524; DB 2; Length 1524;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1524; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTTCAGCGCCTCAATGCGCAGACCTTGAC 60
DB	1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTTCAGCGCCTCAATGCGCAGACCTTGAC 60
QY	61 AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAACCCCAACACGAGAGTTC 120
DB	61 AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAACCCCAACACGAGAGTTC 120
QY	121 GAATGCCGGCGGTCCGCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCT 180
DB	121 GAATGCCGGCGGTCCGCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCT 180
QY	181 TTCTACTCCAATGCTCCCAGGAGATCTTCATCCAGCAGGAAGGGGATACTTTGGGTG 240
DB	181 TTCTACTCCAATGCTCCCAGGAGATCTTCATCCAGCAGGAAGGGGATACTTTGGGTG 240
QY	241 ATATTCCTCGTTGTCCTAGACAATATGAAGAGCCTCACACAAGGTCGTGATCTCAG 300
DB	241 ATATTCCTCGTTGTCCTAGACAATATGAAGAGCCTCACACAAGGTCGTGATCTCAG 300
QY	301 TCCCAAAGACCACCAAGAGCTCTCCAGGAGAAGACCAAGCCAAACGACGAGATAGT 360
DB	301 TCCCAAAGACCACCAAGAGCTCTCCAGGAGAAGACCAAGCCAAACGACGAGATAGT 360
QY	361 CACCAGAAGGTGCACCGTTTTGCATGAGGGTGATCTCATTTGCAAGTTCCACCGGTGCT 420
DB	361 CACCAGAAGGTGCACCGTTTTGCATGAGGGTGATCTCATTTGCAAGTTCCACCGGTGCT 420
QY	421 TTCGTGCTCTCAACGACCACCACTGATGTGTTGTTCTTCTTCTTACTGACACCAAC 480

Db	1501	CAGTCTCCGAGGGGTGTGGCTTAA	1524	
RESULT 3				
ABX70612				
ID	ABX70612	standard; cDNA; 1524 BP.		
AC	ABX70612;			
XX				
DT	26-MAR-2003	(first entry)		
XX				
DE	Peanut Ara h3 cDNA sequence.			
XX				
KW	Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; IgE binding site;			
KW	anaphylactic food allergen; antiallergenic; vaccine; wound healing.			
XX				
OS	Arachis hypogaea.			
XX				
PN	WO200274250-A2.			
XX				
PD	26-SEP-2002.			
XX				
PF	18-MAR-2002; 2002WO-US0009108.			
XX				
PR	16-MAR-2001; 2001US-0276822P.			
PR	18-MAR-2002; 2002US-00276822.			
XX				
XX	(PANA-) PANACEA PHARM.			
PA				
XX				
PI	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;			
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;			
PI	Rabjohn PA, Shin DS, Stanley JS;			
XX				
DR	WPI; 2003-018765/01.			
DR	P-PSDB; ABUS2484.			
XX				
PT	New modified anaphylactic food allergen, useful for preventing or			
PT	treating allergic reactions associated with e.g. anaphylactic allergens.			
XX				
PS	Example 16; Fig 68A; 300pp; English.			
XX				
CC	The invention relates to a modified anaphylactic food allergen has an			
CC	amino acid sequence that is substantially identical to that of natural			
CC	anaphylactic food allergen, except for a cysteine residue that has been			
CC	modified so that it cannot participate in the disulphide bond. The			
CC	modification may also comprise mutation of the IgE binding sites to			
CC	reduce allergenicity. Also included are: (1) a method of making a			
CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding			
CC	or for causing a site specific mutation in the modified anaphylactic food			
CC	allergen; (3) a transgenic plant or animal expressing the modified			
CC	anaphylactic food allergen; (4) a method of treating an individual by			
CC	reducing the clinical response to a natural anaphylactic food allergen;			
CC	and an isolated fragment of peanut allergen Ara h 1. The modified			
CC	anaphylactic food allergen is useful for preventing or treating allergic			
CC	reactions associated with any natural allergen such as food, insect,			
CC	rubber or preferably anaphylactic allergens. It is also useful for			
CC	treating wounds in mammals such as bovine, canine, feline, caprine,			
CC	ovine, porcine, murine or equine species. The present sequence is a cDNA			
CC	encoding a peanut allergen (e.g. Ara h1, h2 or h3)			
XX				
SQ	Sequence 1524 BP; 455 A; 390 C; 398 G; 281 T; 0 U; 0 Other;			
	Query Match	100.0%;	Score 1524;	DB 8; Length 1524;
	Best Local Similarity	100.0%;	Pred. No. 0;	
	Matches 1524; Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	CGGAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60	
Db	1	CGGAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60	
QY	61	AATCGCATTAATCAGAGGGCGGTGTACATTGAGACTTGGACCCCAACACGAGGATTC	120	
Db	61	AATCGCATTAATCAGAGGGCGGTGTACATTGAGACTTGGACCCCAACACGAGGATTC	120	

QY	121	GAATGCGCGGGTGGCCCTCTCTCGCTTAGTCTCTCGCCGCAAGCCCTTGGTAGSCCT	180	
Db	121	GAATGCGCGGGTGGCCCTCTCTCGCTTAGTCTCTCGCCGCAAGCCCTTGGTAGSCCT	180	
QY	181	TTCTACTCCAATGCTCCCGAGGAGATCTTTCAATCCAGCAAGGAGGATACATTTGGGTTG	240	
Db	181	TTCTACTCCAATGCTCCCGAGGAGATCTTTCAATCCAGCAAGGAGGATACATTTGGGTTG	240	
QY	241	ATATTCCCTGGTGTCTTAGACACTATGAAGAGCCTCACACACAAGGTCGTGATCTCAG	300	
Db	241	ATATTCCCTGGTGTCTTAGACACTATGAAGAGCCTCACACACAAGGTCGTGATCTCAG	300	
QY	301	TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGAACCAAAAGCCAAACAGCAACGAGATAGT	360	
Db	301	TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGAACCAAAAGCCAAACAGCAACGAGATAGT	360	
QY	361	CACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCAATTTGGCTGGGAAACACGGAG	420	
Db	361	CACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCAATTTGGCTGGGAAACACGGAG	420	
QY	421	TTCTGGCTCTACAAGCAGCAGCAGCACTGATGTTGTTGCTGTTTCTTCTTACTGACACCAAC	480	
Db	421	TTCTGGCTCTACAAGCAGCAGCAGCACTGATGTTGTTGCTGTTTCTTCTTACTGACACCAAC	480	
QY	481	AACAACGCAACACGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG	540	
Db	481	AACAACGCAACACGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG	540	
QY	541	CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT	600	
Db	541	CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT	600	
QY	601	AGCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGAGCGTGAATTTAGCCCTCGAGA	660	
Db	601	AGCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGAGCGTGAATTTAGCCCTCGAGA	660	
QY	661	CAGCAGCCGCGAGAGAACGAGCAGGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720	
Db	661	CAGCAGCCGCGAGAGAACGAGCAGGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720	
QY	721	AGCGGTTTCAAGAGTTCCTTGGAAACAAAGCTTCCAGGTTTCCAGGTTTCCAGGTTTCCAGG	780	
Db	721	AGCGGTTTCAAGAGTTCCTTGGAAACAAAGCTTCCAGGTTTCCAGGTTTCCAGGTTTCCAGG	780	
QY	781	CAAAACCTTAAGAGGCGAGACCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840	
Db	781	CAAAACCTTAAGAGGCGAGACCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840	
QY	841	GGCCTCAGAAATCTTGAAG	900	
Db	841	GGCCTCAGAAATCTTGAAG	900	
QY	901	GAAATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGA	960	
Db	901	GAAATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAATATGA	960	
QY	961	GGGAATGGTATTTGAAG	1020	
Db	961	GGGAATGGTATTTGAAG	1020	
QY	1021	AGATCCCTTGACATCTTACAACCCCTCAAGCTGGTTCACCTCAAAACCTGCGCAACGATCTCAAC	1080	
Db	1021	AGATCCCTTGACATCTTACAACCCCTCAAGCTGGTTCACCTCAAAACCTGCGCAACGATCTCAAC	1080	
QY	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGAATATGAATATGAATATGAATATGA	1140	
Db	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTGTGAATATGAATATGAATATGAATATGAATATGA	1140	
QY	1141	TTGTTTGTGGCTCACTTACAACACCAAGCAGCATCATATATCGATTGAGGGGACGG	1200	
Db	1141	TTGTTTGTGGCTCACTTACAACACCAAGCAGCATCATATATCGATTGAGGGGACGG	1200	

QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTGTCGATCTCAG 300
Db ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTGTCGATCTCAG 300
QY 301 TCCCAAGAGACCAAGAGAGTCTCCAAAGGAGAGACCAAAAGCCAAAGAGAGATGT 360
Db TCCCAAGAGACCAAGAGAGTCTCCAAAGGAGAGACCAAAAGCCAAAGAGAGATGT 360
QY 361 CACCAGAGGTGCACCGTGTTCGATGAGGAGTCTCAATTCGAGTTCCTCCACCGGTGTGCT 420
Db CACCAGAGGTGCACCGTGTTCGATGAGGAGTCTCAATTCGAGTTCCTCCACCGGTGTGCT 420
QY 421 TTCTGGCTCTACACGACACGACACTGATGTGTGTGCTGTCTCTTACTGACCAAC 480
Db TTCTGGCTCTACACGACACGACACTGATGTGTGTGCTGTCTCTTACTGACCAAC 480
QY 481 AACACGACAAACCGTGTGATCTAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGGAG 540
Db AACACGACAAACCGTGTGATCTAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGGAG 540
QY 541 CAAGAGTCTTAAGGTACACGACCAACAGCAGACAAAGCAGACGAAGAGCTTTACCATAT 600
Db CAAGAGTCTTAAGGTACACGACCAACAGCAGACAAAGCAGACGAAGAGCTTTACCATAT 600
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
QY 661 CAGCAGAGCCGAGAGAACGAGCAGGACGACGAGAGAGAAACGAAAGTGGAAACATCTTC 720
Db CAGCAGAGCCGAGAGAACGAGCAGGACGAGAGAGAGAAACGAAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGCACGACACAGATAGTG 780
Db AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGCACGACACAGATAGTG 780
QY 781 CAAAACTTAAGAGCGGAGACCGAGAGTGAAGAGAGAGGAGCCATTTGACAGTGAGGGGA 840
Db CAAAACTTAAGAGCGGAGACCGAGAGTGAAGAGAGAGGAGCCATTTGACAGTGAGGGGA 840
QY 841 GGCTCAGATCTTGAGCCGAGATAGAAAGAGAGAGTGCAGAAAGAGGAATACGAT 900
Db GGCTCAGATCTTGAGCCGAGATAGAAAGAGAGAGTGCAGAAAGAGGAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGATAGAAAGCGGTGGCAGGGAAGCAGAGCAGG 960
Db GAAGATGAATATGAATACGATGAAGAGATAGAAAGCGGTGGCAGGGAAGCAGAGCAGG 960
QY 961 GGGAAATGGTATGAAGAGACGATCTGCACCGCAAGTGTCTAAAAGAAACATTTGGTAGAAC 1020
Db GGGAAATGGTATGAAGAGACGATCTGCACCGCAAGTGTCTAAAAGAAACATTTGGTAGAAC 1020
QY 1021 AGATCCCTGACATCTACACCTCTCAGCTGGTTCCTCAAACTGCCAACGATCTCAAC 1080
Db AGATCCCTGACATCTACACCTCTCAGCTGGTTCCTCAAACTGCCAACGATCTCAAC 1080
QY 1081 CTTCTAATACCTTAGTGGCTTGACCTAGTGTGATATGAAATCTCTACAGGAATGCA 1140
Db CTTCTAATACCTTAGTGGCTTGACCTAGTGTGATATGAAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db TTGTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
QY 1201 GCTCAGGTGAAGTCTGGGACGACCAAGCGCAACAGAGTGTACACGAGGCTTCAAGAG 1260
Db GCTCAGGTGAAGTCTGGGACGACCAAGCGCAACAGAGTGTACACGAGGCTTCAAGAG 1260
QY 1261 GGTCAAGTGTGTGGTGCCACAGAACTTCGCGGTCTGCTGGAAGTCCACGAGCGAGAAC 1320
Db GGTCAAGTGTGTGGTGCCACAGAACTTCGCGGTCTGCTGGAAGTCCACGAGCGAGAAC 1320
QY 1321 TTCGAATACGTGGCATTCAAAGACAGACTCAAGGCCCCAGCATAGCCAACTTCGCGGTGAA 1380

Db 1321 TTCGAATACGTGGCATTCAAAGACAGACTCAAGGCCCAGCATAGCCAACTTCGCGGTGAA 1380
QY 1381 AACTCCGTCTATAGATAAAGTCCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 1381 AACTCCGTCTATAGATAAAGTCCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
QY 1441 GAGCAGGCAAGGAGCTTAAGAACAAACCCCTTCAAGTTCCTCGTTCACCGTCTCAG 1500
Db 1441 GAGCAGGCAAGGAGCTTAAGAACAAACCCCTTCAAGTTCCTCGTTCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
RESULT 5
AAS08540
ID AAS08540 standard; cDNA; 1524 BP.
XX
AC AAS08540;
XX DT 23-OCT-2001 (first entry)
XX DNA encoding anaphylactic antigen Ara h 3.
XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
XX allergy; mast cell; basophil; mouse; ss.
XX Mus sp.
FH Key Location/Qualifiers
FT 1. .1533
FT /*tag= a
FT /product= "Peptide antigen Ara h 3"
FT /note= "Coding sequence of amino acids 1-3 not given"
XX WO200140264-A2.
XX 07-JUN-2001.
XX 06-DEC-2000; 2000WO-US033124.
XX 06-DEC-1999; 99US-00455294.
PR 23-JUN-2000; 2000US-0213765P.
PR 27-SEP-2000; 2000US-0235797P.
XX (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
DR P-PSDB; AAU04708.
XX Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E.
XX Disclosure; Fig 11; 100pp; English.
XX The sequence represents the coding sequence of anaphylactic antigen Ara h 3. Ara h 3 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind immunoglobulin E (IgE) as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who

CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 1524 BP; 455 A; 392 C; 396 G; 281 T; 0 U; 0 Other;

Query Match 99.8%; Score 1520.8; DB 4; Length 1524;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGG CAG CAA CCG GAG GAG AAG CGG TGC CAG TTC CAG CGC CTCAAT GCG CAG ACG CTTC GAC 60
Db 1 CGG CAG CAA CCG GAG GAG AAG CGG TGC CAG TTC CAG CGC CTCAAT GCG CAG ACG CTTC GAC 60

Qy 61 AAT CGC ATT GAA TCAG AAG GCG GTT ACATT GAG ACT TTG GAA CCCC CCA CCA CCG AGG ATT C 120
Db 61 AAT CGC ATT GAA TCAG AAG GCG GTT ACATT GAG ACT TTG GAA CCCC CCA CCA CCG AGG ATT C 120

Qy 121 GAAT GCG CGG CGT GCG CCT CTCT CGCT TAGT CTT CCG CGC GAA CCG CCT TCG TAG GCCT 180
Db 121 GAAT GCG CGG CGT GCG CCT CTCT CGCT TAGT CTT CCG CGC GAA CCG CCT TCG TAG GCCT 180

Qy 181 TTCT ACTT CCAAT GCT CCCC AGG AGAT CTT CAT CAG CAA GAA GGG GAT CTT TGG GTTG 240
Db 181 TTCT ACTT CCAAT GCT CCCC AGG AGAT CTT CAT CAG CAA GAA GGG GAT CTT TGG GTTG 240

Qy 241 ATATT CCTT GGT TGT CTCTAG ACATAT GAA GAG CTT CAC A CAG AGT TCG ATCT CAG 300
Db 241 ATATT CCTT GGT TGT CTCTAG ACATAT GAA GAG CTT CAC A CAG AGT TCG ATCT CAG 300

Qy 301 TCC CAA AGA CCA CCA GAG CTT CCA AGG AAG AAG CAA AGC CAA CAG CAG ACG AGT 360
Db 301 TCC CAA AGA CCA CCA GAG CTT CCA AGG AAG AAG CAA AGC CAA CAG CAG ACG AGT 360

Qy 361 CAC CAG AAG GTG CAC CGT TTT CGAT GAG GTT GAT CTT CAT TGC AGT TCC CAC CGT TGT GCT 420
Db 361 CAC CAG AAG GTG CAC CGT TTT CGAT GAG GTT GAT CTT CAT TGC AGT TCC CAC CGT TGT GCT 420

Qy 421 TTT CTT GGT CTT CAC ACG CAG CAG CACT GAT GTT GGT CTT TCT TCT TACT GAC CCA C 480
Db 421 TTT CTT GGT CTT CAC ACG CAG CAG CACT GAT GTT GGT CTT TCT TCT TACT GAC CCA C 480

Qy 481 AAC ACG CAA CCA CCG TTT GAT CAG TTT CCG CAG GAG ATT CAA TTT GGT GGG AAG CAG CAG 540
Db 481 AAC ACG CAA CCA CCG TTT GAT CAG TTT CCG CAG GAG ATT CAA TTT GGT GGG AAG CAG CAG 540

Qy 541 CAAG AGT TTT TAA GGT TACC AGC CAA CAA GAG CAG CAA AGC CAA GAG CTT ACC ATAT 600
Db 541 CAAG AGT TTT TAA GGT TACC AGC CAA CAA GAG CAG CAA AGC CAA GAG CTT ACC ATAT 600

Qy 601 AGCC CATA CAG CCG CAA AGT CAG CTTAG CAG AAG AGC GTG AAT TTTAG CCG CTT CAG GGA 660
Db 601 AGCC CATA CAG CCG CAA AGT CAG CTTAG CAG AAG AGC GTG AAT TTTAG CCG CTT CAG GGA 660

Qy 661 CAG CAG CAG CCG CAG AAG CAG CAG GAG CAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 720
Db 661 CAG CAG CAG CCG CAG AAG CAG CAG GAG CAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 720

Qy 721 AGCG GCTT CAC GCG GAG TTT CTT GGA A CAA GCG CTT CAG GTT GAC CAG CAG CAG ATAG TG 780
Db 721 AGCG GCTT CAC GCG GAG TTT CTT GGA A CAA GCG CTT CAG GTT GAC CAG CAG CAG ATAG TG 780

Qy 781 CAAA A CTTAAG CCG CAG ACG CAG AGT GAA AAG AGG AGC CATT TGT GAC AGT GAG GGA 840
Db 781 CAAA A CTTAAG CCG CAG ACG CAG AGT GAA AAG AGG AGC CATT TGT GAC AGT GAG GGA 840

Qy 841 GGC CTT CAG AAT CTT GAG CCG CAG ATAG AAG AAG AAG CAG CCG CCG CAG AAG AAG AAG AAG AAG 900
Db 841 GGC CTT CAG AAT CTT GAG CCG CAG ATAG AAG AAG AAG CAG CCG CCG CAG AAG AAG AAG AAG AAG 900

Db 841 GGC CTT CAG AAT CTT GAG CCG CAG ATAG AAG AAG AAG CCG CCG CAG AAG AAG AAG AAG AAG 900
Qy 901 GAAG ATGA ATAT GAAT CAG ATGA AGG ATAG AAG CCG TGG CAG GGA AGC AAG CAG AGC AGG 960
Db 901 GAAG ATGA ATAT GAAT CAG ATGA AGG ATAG AAG CCG TGG CAG GGA AGC AAG CAG AGC AGG 960
Qy 961 GGA AAT GGT ATT GAA GAG ACG ATCT CCA CCG CAG GGT GCT TAA AAG AAG AAG AAG AAG AAG 1020
Db 961 GGA AAT GGT ATT GAA GAG ACG ATCT CCA CCG CAG GGT GCT TAA AAG AAG AAG AAG AAG AAG 1020

Qy 1021 AGAT CCCC CTG ACAT CTTACA AACCCT CAAG CTGG TTTCACT CAAA AACTG CCAAC GATCTCTCAAC 1080
Db 1021 AGAT CCCC CTG ACAT CTTACA AACCCT CAAG CTGG TTTCACT CAAA AACTG CCAAC GATCTCTCAAC 1080

Qy 1081 CTTCTA ATACTT TAG GTGGCTTGGACCTTAGTCTGCTGAATATGGAATATCTCTACAGAAATGCA 1140
Db 1081 CTTCTA ATACTT TAG GTGGCTTGGACCTTAGTCTGCTGAATATGGAATATCTCTACAGAAATGCA 1140

Qy 1141 TTGTTTGT CGCTCACTACACCAACG CACACAGCATCATATATCGATTGAGGGACGG 1200
Db 1141 TTGTTTGT CGCTCACTACACCAACG CACACAGCATCATATATCGATTGAGGGACGG 1200

Qy 1201 GCTCAG CTGCAAGTCTGGACAGCAACG CCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAG CTGCAAGTCTGGACAGCAACG CCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260

Qy 1261 GGTCA CGTGTCTGTGTG CCAAGAACTTCGCGCTCGCTGGAAAGTCCCAAGACGAGAAC 1320
Db 1261 GGTCA CGTGTCTGTGTG CCAAGAACTTCGCGCTCGCTGGAAAGTCCCAAGACGAGAAC 1320

Qy 1321 TTGCA ATACTGTCATTTCAAGCAGACTCAAGGCC CAGCATAGCCAACTCCGCGGTGAA 1380
Db 1321 TTGCA ATACTGTCATTTCAAGCAGACTCAAGGCC CAGCATAGCCAACTCCGCGGTGAA 1380

Qy 1381 AACTCCGT CATATAGATAA CCTCGCGAGGAGTGTGTCAAAATTCATATGSCCTCCAAAGG 1440
Db 1381 AACTCCGT CATATAGATAA CCTCGCGAGGAGTGTGTCAAAATTCATATGSCCTCCAAAGG 1440

Qy 1441 GAGCAG CAGGAGCAGCTTAAGAA CAAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Db 1441 GAGCAG CAGGAGCAGCTTAAGAA CAAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500

Qy 1501 CAGTCTCCAGGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCAGGGCTGTGGCTTAA 1524

RESULT 6
AAF90338
ID AAF90338 standard; cDNA; 1855 BP.
XX
AC AAF90338;
XX AC
XX 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX
XX Peanut allergen Ara h4 cDNA.
XX Peanut; allergen; Ara h2; Ara h3; transgenic plant; allergy; ss.
XX Arachis hypogaea.
OS
FH Key Location/Qualifiers
CDS 1..1593
FT /*tag= a
FT misc_feature 430..1190
FT /*tag= b
FT /note= "PCR amplified fragment"
XX
XX W0200136621-A2.
XX
XX 25-MAY-2001.
XX

[illegible]

Db 1127 TGGTACAAAGTGGTCAATTCGATGAGAGAGTGTGGTGGAGAGCTGCAGAGGGAC 1186
 Qy 1265 ACATGCTTGTGGTCCAGCAACTTCGCGCTGCTGGAAGTCCAGAGGAGAACTTCG 1324
 Db 1187 AGGTGTTAAATGTGCCCAAACTTTGCGGTGGCTGCAAGATCACAGAGCGCAACTTCG 1246
 Qy 1325 AATACGTGGCATTCAGACAGACTCAAGGCCAGCTAGCCACCTGCGCGGTGAAACT 1384
 Db 1247 AGTATGTTTCATTCAGACCAATGATAGACCTCGTGGCAACTTCGAGGTGCAAACT 1306
 Qy 1385 CCGTCATAGATAAATCTCCGAGAGAGTGGTGGCAAAATTCATATGGCTCCAAAGGGAGC 1444
 Db 1307 CATTTGTTGAAGCATTCGCGGAGAGTGTTCAGCAAACTTTAACTAAGGAGGAGC 1366
 Qy 1445 AGCAAGGAGCTTAAGAAACAAACCCCTTCAAGTTCCTGTTCCACCGTCTCAGCAGT 1504
 Db 1367 AGGCAGGAGCTCAAGAAACAAACCCCTTCAAGTTCCTGTTCCACCTAAGGAGTCTC 1426
 Qy 1505 CTCCGAGGCTGTGGCTTA 1523
 Db 1427 AGAGGAGTGTGGCTTA 1445

RESULT 9

ADH89252
 ID ADH89252 standard; DNA; 1446 BP.

XX AC ADH89252;

XX XX 06-MAY-2004 (first entry)

XX DT G. max glycinein subunit G3 DNA.

XX DE ds; gene; double stranded RNA; storage protein; 2S-albumen; 7S-globulin;
 XX KW 11S/12S-globulin; zein-prolamine; homogenistate metabolic pathway;
 KW pharmaceutical; plant; abiotic stress; fatty acid composition;
 KW lipid composition; oil composition; carbohydrate composition; colour;
 KW pigmentation; pathogen resistance; fruit ripening delay; aging;
 KW male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;
 KW caffeine; theophylline; threonine biosynthesis; glycinein.
 XX OS Glycine max.

XX PN WO2003078629-A1.

XX XX 25-SEP-2003.

XX XX 17-MAR-2003; 2003WO-EP002735.

XX XX 20-MAR-2002; 2002DE-01012892.

XX XX (BADI) BASF PLANT SCI GMBH.

XX XX Kock M, Bauer J;

XX XX WPI; 2003-803889/75.

DR P-PSDB; ADH89253.

XX XX Reducing expression of at least two target genes, useful e.g. for
 producing transgenic plants, using partly double-stranded interfering
 RNA.

XX PS Claim 7; SEQ ID NO 27; 228pp; German.

XX CC This invention describes a novel method for reducing the expression of at
 CC least two different endogenous target genes in a eukaryotic cell or
 CC organism by introducing an RNA molecule that is at least partly double
 CC stranded. The transcribed RNAs from at least two target genes have
 CC homology below 90% and the RNA molecule is formed as a single, self-
 CC complementary molecule. At least one of the double-stranded structures
 CC formed from individual sense sequences has an even number of repeats of
 CC 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At

CC least two target genes are selected from different classes of storage
 CC protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-
 CC prolamine and at least one of the sense sequences is identical to storage
 CC protein sequences or genes in the homogenistate metabolic pathway or
 CC enzyme types, e.g. acetyl transferases, thioesterases, (de)branching
 CC enzymes or cellulases. The RNA of the invention, also related cassettes,
 CC expression systems, vectors and transgenic organisms are used for
 CC preparation of pharmaceuticals, in biotechnological processes and plant
 CC biotechnology, specifically in plants to improve protection against
 CC abiotic stress, to modify composition and/or content of fatty acids,
 CC lipids and oils, to modify carbohydrate composition, to alter colour or
 CC pigmentation, to reduce content of storage proteins, to increase
 CC resistance to pathogens, to inhibit stem break, to delay fruit ripening
 CC or aging, to induce male sterility, to reduce content of toxic or
 CC unwanted components, to modify lignification and/or lignin content, to
 CC modify the fibre component in foods or fibre quality in cotton, to reduce
 CC susceptibility to shock, to increase synthesis of Vitamin E, to reduce
 CC contents of nicotine, caffeine or theophylline and to increase methionine
 CC content, by reducing threonine biosynthesis. The method provides a rapid
 CC and efficient way of reducing gene expression, can inhibit more than one
 CC target gene, prevents development of multiple phenotypes (since the
 CC transcription rate is the same for all RNA sequences, significantly
 CC reducing the selection process required to produce an organism with
 CC effective suppression of all target genes), avoids problems of epigenic
 CC gene silencing, does not require synthesis of individual RNA sequences
 CC and the method can be applied to plants with complex (polyploid) genomes.
 CC No interference between the individual RNA sequences occur. This sequence
 CC represents a target gene used in the method of the invention.
 XX XX
 SQ Sequence 1446 BP; 426 A; 361 C; 354 G; 305 T; 0 U; 0 Other;

Query Match 34.0%; Score 517.4; DB 10; Length 1446;

Best Local Similarity 62.8%; Pred. No. 1.4e-137;

Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

Qy 5 AGCAACCGGAGGAGAACGGCTGCGAGTCCAGCGCTCAATGGCGAGAGACCTGACAAATC 64
 Db 71 AGCAGCCACAGCAAAACGAGTGCCAGATCCAGCGCTCAATGGCTTAAACCGGATTAACC 130
 Qy 65 GCATTCAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACACAGGAGTTCGAAT 124
 Db 131 GTATAGAGTCAGAAAGGTGGCTTCATTGAGACATGGAACCTCAACAACAGCCATTCCAGT 190
 Qy 125 GCGCCGGCTGCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCGCTTCT 184
 Db 191 GTGCGGTGTGCGCTCTCTCGCTGCAACCCCTCAACCGCAACGCCCTTCGCAGACCTTCT 250
 Qy 185 ACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGATCTTTGGGTGATAT 244
 Db 251 ACACCAACGCTCCCGAGGAGATCTATCCACAGGTAGTGTATTTTGGCATGATAT 310
 Qy 245 TCCCTGGTTGTCTTAGACACTATGAAGAGCTTCACACACAAGGTCTCGATCTCAGTCCC 304
 Db 311 TCCCGGTGTCTTAGACACTTTGAAGAGCT----- 342
 Qy 305 AAGACCAACAGAGCGTCTCCAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGTACC 364
 Db 343 -----CAACAAAAGGACAAAAGCAGCAGCCCAAGACCGTCAAC 382
 Qy 365 AGAAGGTGCACCGTTTCGATGAGGGTGATCTCAATTCAGTTCCACCGGTGTTGCTTCT 424
 Db 383 AGAAGATCTATCTTCAGAGAGGGTGATTTGATTCAGTGCACCGGTGTTGCTTCTACT 442
 Qy 425 GGCTCTACAACGACCACTGATGTTGTTGCTGTTCTTCTTACTGACCAACAACA 484
 Db 443 GGATGTACAACATGAAGACACTCTCTGTTGTCCTTTCTTTATTGACCAACAGCT 502
 Qy 485 AGCAACACAGCTTGTATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAG 544
 Db 503 TCCAGAACACAGCTCGACAGATGCTCTAGGAGATTTCTATCTTGTGGGAACCAAGCAAG 562
 Qy 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 604

251	Db	ACACCAACGCTCCCCAGGAGATCTCATCCAAAGGTAGTGGTATTTTTTGGCANGATAT	310
245	Qy	TCCCTGGTGTCTTAGACA CATATGAAGAGCCTCAACAAGGTGCTCGATCTCAGTCCC	304
311	Db	TT	342
305	Qy	AAAGACCAACGAAGAGCTCTCAAGGAGAGAACCAAGCAACAGCAAGATATGTCACC	364
343	Db	-----CAACAAAAGGACAAAGCAGCAGGCCCCCAAGACCGTCAACC	382
365	Qy	AGAAAGTGCACCGTTTTCGATCAGGGTGATCTCATTTGCAAGTTCCACACGGTGCTGTTCT	424
383	Db	AGAAAGATCTATCACTTCAGAGGGGTGATTTGATTCAGTGCCAAACCGTTTTCGATCT	442
425	Qy	GGCTCTCAACGACCAACGACACTGATGTGTGTGCTGTTCTTCTTACTGACACCAACAACA	484
443	Db	GGATGTACAACAATGAAGACACTCCTGTTGTTGCGGTTTCTTTATTGACACCAACAGCT	502
485	Qy	ACGACAACCAAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAAACAGGAGCAAG	544
503	Db	TCCAGAACCAGCTCGACAGATGCTTAGGAGATTCTATCTTTGCTGGGAACCAAGAGCAAG	562
545	Qy	AGTTCTTAAAGTACACAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTTACCATATAGCC	604
563	Db	AGTTTCTACAGTATCAGCCACAGACGACGAG-----	595
605	Qy	CATACAGCCGCAAAAGTCAGCTTAGACAAGAAGAGCGTGAATTTAGCCCTCGAGGACAGC	664
596	Db	-----GAGGTACTC	604
665	Qy	ACAGCCGAGAGAACGAGCAGGACAAGAAGAAGAAAACGAAGGTGGAAAACATCTTCAGCG	724
605	Db	AAAAAGCAAGGAAAAGCGCTCAGCAAGAAGAAGAAAACGAAGGAGGAGCATATTGAGTG	664
725	Qy	GCTTCAGCCGCGAGTTCTTGAAACAAAGCTTCAGAGTTGACGACAGACAGATAGTGCAAA	784
665	Db	GCTTCGCCCCGGAATCTTGAAACATGCGTTCT---GTGCTGGACAGGACAGATAGTGAAG	721
785	Qy	ACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGGAGCGCATTTGTCACAGTGAAGGGAGGCC	844
722	Db	AGCTACAAGGTGAGAACGAAGAGAGAGAGAGGGTGCCATTTGTGACAGTGAAGAGGATC	781
845	Qy	TCAGAACTTTGAGCCCGAGATAGAAGAGACGTGCGCAGCAAGAGAGAGGAATACGATGAAG	904
782	Db	TCAGCGTGATAAGGCCCAACCCACGGAAGCAGCAGCAACAAAGACCCGAGAGAGGAGAAGC	841
905	Qy	ATGAATATGATACGATGACAGAGATAGAGGCGTGGCAGGGGAAGCAGAGGACGGGGGA	964
842	Db	CAGATTGTGACGAGAAAGACAAACAT-----TGCCAAAGCCAAAGACGAGAA	886
965	Qy	ATGGTATTGAAGAGACGATCTGACCGCAGTGTCTAAAGAAACATTTGGTAGAAACAGAT	1024
887	Db	ATGSCATTTGACGAGACCATTTTGCACAATGAGACTTCGCCCAACCATTTGGCCAGACTTCAT	946
1025	Qy	CCCTGACATCTCAACCCCTCAAGCTGGTTCACTCAAAACTGCCAACGATCTCAACCTTC	1084
947	Db	CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACACCGCTACCAGCGCTCGACTTCC	1006
1085	Qy	TAATACTTAGTGGCTTGGACCTAGTCTGATATGGAATCTCTACAGGATGATCTGT	1144
1007	Db	CAGCCCTCTCGTGCTCAAACTCAGTGCCCAAGTTTGGATCATCTCCGCAAGATGCTATGT	1066
1145	Qy	TTGTGCTGCTACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC	1204
1067	Db	TCGTGCCACACTACAACCTGAACCGAAACAGCATTAATATACGCAATTTGAATGGACGGGCAT	1126
1205	Qy	ACGTGCAAGTCGTGGACAGCAACGGCAACAGAGGTGACGACGAGGAGCTTCAAGAGGGTC	1264
1127	Db	TGGTACAAGTGTGAATTGCAATTGTCAGAGAGAGTGTGATGGAGAGCTGCAAGAGGGAC	1186
1265	Qy	ACGTGCTTTGGTGGCAGAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAACTTCG	1324
1187	Db	AGGTGTTAAATTGGTCCCAAACTTTTGGGTGGCTGCAAGATCAACAGAGCAACTTCG	1246

RESIT.T 11

RESUL TI
AAV17565

AAV17565
ID AAV17565 standard; CDNA; 1488 bp.

XX
XX

AC AAV17565;

XX

DT 10-JUN-1998 (first entry)

XX

DE Coding sequence for Gyl subunit of glycinin.

Beta-conglycinin: soybean seed

KW Beta-conglyc
KW seed storageXX
MY
seed storage

Glycine max.

XX

PN WO9747731-A2

	1
	6
	9
	7
	8
	3
	4
	5
	2
X	1
X	6
X	9
X	7
X	8
X	3
X	4
X	5
X	2

PD 18-DEC-1997.

XX 10-JUN-1997 07WOC-

PF 166T-NUN-OT 9/WO-XX

PR 14-JUN-1996: 96US-0019940P.

XX
XX
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840

(DUPO) DU PONT DE NEMOURS & CO E I.
Kinney AJ, Fader GW;
WPI; 1998-052298/05.
Suppression of specific classes of soybean seed protein genes - useful to change seed storage protein profiles of transgenic plants.

XX

PS Disclosure; Page 34-35; 58pp; English.

Query Match		33.7%;	Score 513;	DB 2;	Length 1488;
Best Local Similarity		62.7%;	Pred. No. 2.6e-136;		
Matches		968;	Conservative	0;	Mismatches 425; Indels 150; Gaps 5;
QY	5	AGCAACGGAGGAAACGCGTGCAGTTCAGCGCCCTCAATGCGCAGAGACCTGACAATC	64		
DB	71	AGCAGCCTCAGCAAAACGAGTGCAGATCAAAAACCTCAATGCGCTCAAAACCGGATAACC	130		
QY	65	GCATTGAATCAGAGGGCGTTACATTGAGACTTTGGAAACCCCAACACACCGAGTTCGAAT	124		
DB	131	GTATAGAGTCAGAGGAGGGCTCATTTGAGACATGGAAACCCCAACACCGATTCAGT	190		
QY	125	GCGCCGCGTGCCTCTCTCGCTTAGTCTCCGCGCGCAACGCGCTTCGTAGCGCTTCT	184		
DB	191	GTGCGGCTGTGCGCTCTCTCGCTGACCTCAACCGCAACGCGCTTCGTAGACCTTCT	250		
QY	185	ACTCCAATGTCCTCCAGGAGATCTTCATCCAGAAGGAGGATATCTTTGGTTGATAT	244		
DB	251	ACACCAACGGTCCCAGGAATCTACATCCAACAAGGTAAAGGTATTTTGGCATGATAT	310		
QY	245	TCCTGTTGCTTAGACATATGAAGAGCCTCACACACAGGTGCTCGATCTCAGTCCC	304		
DB	311	ACCGGGTTCCTAGCACATTTGAAGAGCTCA-----	344		
QY	305	AAAGACCACCAAGACGTCTCCAAGGAGAGAACCAAGCCAAAGCAGACGAGATAGTCACC	364		
DB	345	-----ACAACCTCAACAAGAGGACAAGAGCAGACACACAGACCGTCAACC	391		
QY	365	AGAAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCGATTCCTCCCGGTGTTGCTTCT	424		
DB	392	AGAAGATCTATACTTCAGAGAGGGTGATTTGATCGCAGTGCCTACTGTGGTTGTCATGTT	451		
QY	425	GGCTCTACACGACCGACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACA	484		
DB	452	GSATGTACAACAATGAAGACACTCCTGTTGTTGCGGTTTCTATTATTGACACCAACAGT	511		
QY	485	ACGACAACACGCTTGATTCAGTTCCTCCAGGAGATTCAATTTGGCTGGCAACCGGCAAG	544		
DB	512	TGGAACACGACTCGACCATGCTTAGGATTTCTATCTTGTGGGAACCAAGACGAG	571		
QY	545	AGTTCTTAAGGTACAGACAAACAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC	604		
DB	572	AGTTCTTAATAATTCAGCAAG-----	594		
QY	605	CATACAGCCCGCAAGTCAAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGACAGC	664		
DB	595	-----	607		
QY	665	ACAGCCGACAGAACGACGACGAGCAAGAAGAAACGAAGGTGGAAACATCTTCAGCG	724		
DB	608	AAAGCCAGAAAGGAAGCATCAGCAGAGAGAGAAACGAGGAGCGATATTGAGTG	667		
QY	725	GCTTCAGCGCGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGACATAGTCAAA	784		
DB	668	GCTTCACCTGGAATCTTGGAAATGCAATTCAGCT---GGACAAGCAGATAGCGAAAA	724		
QY	785	ACCTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGGAGGGGCC	844		
DB	725	ACCTAAGAGAGAGAACGAAGGGGAAGACAAGGGAGCCATTGTGACAGTGAAGAGGGTC	784		
QY	845	TCAGAAATCTTGAGCCCA-----GATAGAAGAGACGTGCGCGACGAAGAGAGGAAT	895		
DB	785	TGAGCGTGATAAACCACCCACGACGACGACGCAACAAGACCCCGAGAGAGAGAGAG	844		
QY	896	ACGATGAAGATGAATA-----TGAATACGATGAAGAGGATAGAAGCGGTG	940		
DB	845	AAGAAGAGGATGAGAAGGCCACAGTGCAGAGGTTAAAGCAAACTGCAACGCCGCCGAG	904		
QY	941	GCAGGGAGACGAGCGAGGGGGAATGTTATTGAAGAGAGATCTGCACCGCAAGTGCTA	1000		
DB	905	GAAGCCAAAGCAAGCAAGAAATGGCATTCACGAGACCATATGCACCATGAGACTTC	964		

Qy	1001	AAAAGAAACATTGGTAGAAAACAGATCCCTCGATCTTACAACCCCTCAAGCTGGTTCACTCA	1060
Db	965	GCCACAACATTGGCCAGACTTTCATCACCTGACATCTACAACCCCTCAAGCCGGTAGCGTCA	1024
Qy	1061	AAACTGCCAAACATCTCAACCTTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATG	1120
Db	1025	CAACCGCCACCAAGCCTTGACTTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTTG	1084
Qy	1121	GAAATCTCTACAGGAATGCATTGTTTGTGCTCACTACAACACAACGCAACAGCATCA	1180
Db	1085	GATCTCTCCGCAAGAAATGCAATGTTCTGTGCCACACTACAACCTGAACGCGAACAGCATAA	1144
Qy	1181	TATATCGATTGAGGGACGGGCTCACGTGCAAGTCTGTGGACAGCAACGGCAACAGAGTGT	1240
Db	1145	TATACGATTGAATGGACGGGCATTGATACAAGTGTGTGAATTGCAACGGTGAGAGTGT	1204
Qy	1241	ACGACGAGGAGCTTCAAGAGGGTCAAGTCTGTGTTGGTGCCACAGAACTTCGCGCTCGCTG	1300
Db	1205	TTGATGGAGAGCTGCAAGAGGGACGGGTCTGATCGTGCACAAAACCTTTGTGGTGGCTG	1264
Qy	1301	GAAAGTCCAGAGGAGCAACTTTCGAATACGTGCGATTCAAGACAGACTCAAGGCCCAAGCA	1360
Db	1265	CAAGATCACAGAGTGCAACTTTCGAGTATGTGTCAATTCAAGACCAATGATACACCCATGA	1324
Qy	1361	TAGCCAACTCGCGGTGAAAACCTCGTCAATAGATAACCTGCGGAGGAGGTGGTTGCAA	1420
Db	1325	TCGGCACTCTTGAGGGGCAAACTCATTTGTGAACGATTAACAGAGGAAGTGAATTCAGC	1384
Qy	1421	ATTCAATATGGCTTCCAAAGGGAGCAGCAAGGCAAGTCTTAAGAAACAACAACCCCTTCAAGT	1480
Db	1385	ACACTTTCAACCTAAAAAGCCAGCAGGCCAGGCAAGATAAAGAAACAACAACCCCTTTCAAGT	1444
Qy	1481	TCCTCGTTCCACCGTCTCAGCAGTCTCCGAGGCGCTGTGGCTTA	1523
Db	1445	TCCTGGTTCCACCTCAGGAGTCTCAGAAGAGAGCTGTGGCTTA	1487
RESULT 12			
ABSS5191			
ID	ABSS5191 standard; DNA; 1488 BP.		
XX			
AC	ABSS5191;		
XX			
DT	17-DEC-2002 (first entry)		
XX			
DE	Glycine max (Soybean) var. Shiotsurunoko gene #1.		
XX			
KW	Soybean; Glycinin; atomic coordinate data; processability; soya protein; Shiotsurunoko; protein co-ordinate data; gene; ds.		
XX			
OS	Glycine max.		
XX			
PH	Location/Qualifiers		
FT	1..1488		
FT	/*tag= a		
FT	/product= "Glycine max (Soybean) var. Shiotsurunoko		
FT	protein #1"		
XX			
PN	JP2002193996-A.		
XX			
PD	10-JUL-2002.		
XX			
PF	21-DEC-2000; 2000JP-00405097.		
XX			
PR	21-DEC-2000; 2000JP-00405097.		
XX			
PA	(KYOU) UNIV KYOTO.		
XX			
DR	WPI: 2002-685438/74.		
XX			
XX	P-PSDB; ABG71264.		
XX			
PT	Glycinin, beta-conglycinin and proglycinin, their crystal structures, three dimensional coordinates, three dimensional structured and models		

PT	and their uses.		
XX	Disclosure; Page 1264-1266; 1298pp; Japanese.		
PS	The present invention relates to a new Glycinin characterised by the		
XX	atomic coordinate data fully defined in the specification. The structure		
CC	can be used for improving processability of soya protein. The present		
CC	nucleic acid sequence encodes the Glycine max (soybean) var.		
CC	Shirotsurunoko protein #1, as described in the specification		
XX			
SQ	Sequence 1488 BP; 466 A; 368 C; 362 G; 292 T; 0 U; 0 Other;		
	Query Match 33.7%; Score 513; DB 6; Length 1488;		
	Best Local Similarity 62.7%; Pred. No. 2.6e-136;		
	Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;		
QY	5 AGCAACCGGAGGAGGAGCGTCCAGTTCACGCGCTCAATGCGCAGAGACTCACAATC 64	Db 785 TGAGCGTGATAAAACCCACCGGACGAGCAGCAAAAGACCCCGAGGAAGAGAAG 844	
Db	71 AGCAGCCTCAGCAAAACGAGTGCCAGATCCAAAACCTCAATGCCCTCAACCGGATTAAC 130	QY 896 ACATGAAGATGAATA-----TGAATACGATGAAGAGGATAGAAGCGGTG 940	
QY	65 GCATTGAATCAGAGGCGGTTTACATTGAGACTTGGAAACCCCAACAACAGGAGTTGCAAT 124	Db 845 AAGAAGAGATGAGAAGCCACAGTGCAGAGGTAAGACAAACACTGCCAACGCCCGAG 904	
Db	131 GTATAGAGTCAGAGGAGGCGTCAATTGAGACATGGAAACCCCTAAACAAGCCATTCCAGT 190	QY 941 GCAGGGAAGCAGAGCGCAGGGGAATGGTATTGAAGAGACAGTCTGCACCGCAAGTCTTA 1000	
QY	125 GCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGTTTCT 184	Db 905 GAAGCCAAAGCAAAAGCAGAGAAATGGCATTTGACGAGACCATATGCACCATGAGACTTC 964	
Db	191 GTGCGGCTGTGCGCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGTAGACCTTCCT 250	QY 1001 AAAAGAACATTGGTAGAAACAGATCCCTGACATCTACACCCCTCAAGCTGGTTCACCTCA 1060	
QY	185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACCTTTGGGTGATAT 244	Db 965 GCACAAACATTGGCCAGACTTTTCATCACTGACATCTACAAACCTCAAGCCCGTAGCGTCA 1024	
Db	251 ACACCAACGCTCCCGAGGAATCTACATCCAAAGGTAAAGGATATTTTGGCATGATAT 310	QY 1061 AAATGCCAACGATCTCAACCTTTCTAATATCTTAGTGGCTTGAGCCTAGTGTGTAATATG 1120	
QY	245 TCCCTGTTGTCTTAGACATATGAAGAGCCTCACACAAAGGTCTGTGATCTCAGTCCC 304	Db 1025 CAACGCCACCCAGCCCTTGACTTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTTG 1084	
Db	311 ACCCGGTTGTCTTAGCACATTTGAAGAGCCTCA----- 344	QY 1121 GAAATCTCTACAGGAATGCATTTGTTGTGCTCACTACAAACCAACGACACACAGCATCA 1180	
QY	305 AAAGACCACCAAGAGCTCTCCAGGAGAGAACCAAGCCACAGCAGAGATAGTCAAC 364	Db 1085 GATCTCTCCGCAAGAAATGCAATGTTGCTGCCACACTACAACTGAAACGCGAACAGCATAA 1144	
Db	345 -----ACAACCTCAACAAAGAGGACAAAGCAGCAGACCAAGACCGTCAAC 391	QY 1181 TATATCGATTGAGGGACCGGCTCAGTGCAGTCTGTGGACAGCAACGCGCAACAGAGTGT 1240	
QY	365 AGAAGTGTACCGTTTCGATGAGGCTGATCTCATTTGCAAGTTCCACCGGTGTGCTTTCT 424	Db 1145 TATACGCAATTGAAATGGACGGGCATTTGATCAAGTGGTGAATTGCAACGGTGAGAGAGTGT 1204	
Db	392 AGAAGATCTAATCTTCAGAGAGGGTGATTTGATCGCAGTGCCTACTGTGTTGCAATGGT 451	QY 1241 ACACGAGGAGCTTCAAGAGGGTCACTGCTGTGTGGTGGCCACAGAACTTTCGCGTGGCTG 1300	
QY	425 GGCTCTTAAACAGCACGACATGATGTTGTTGCTGTTTCTTCTTACTGACACCAACAACA 484	Db 1205 TTGATGGAGAGCTGCAAGAGGACGGGTGCTGTGATCTGTGCCACAAACTTTTGTGGTGGCTG 1264	
Db	452 GGATGTACAAACATGAAGACACTCTCTGTTGTCGGTTCATATTATTGACACCAACAGCT 511	QY 1301 GAAAGTCCCGAGCGAGAACTTTCGAATACGTGGCAATTAAGAGACAGACTCAAGGCCCGCAGCA 1360	
QY	485 ACACAAACGAGTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAG 544	Db 1265 CAAGATCAGAGTGAACAACCTTCGAGTATGTGTTCATTCAAGACCAATGATACACCATGA 1324	
Db	512 TGGAGAACCCAGCTCGACCAATGCTTAGGAGATTCTATCTTGTCTGGGAACCAAGAGCAAG 571	QY 1361 TAGCCAACTCGCGGTGAAACTCTCGTTCATAGATAAACCTCCGCGAGGAGTGGTTGCAA 1420	
QY	545 AGTTCTTAGGTTACCAACAAAGCAGACAAAGCAGACGAAGCTTACCATATAGCC 604	Db 1325 TCGGCACCTTTGACGGGGCAACTCATTTGTTGAACGCATTAACACAGAGGAAGTGAATTCA 1384	
Db	572 AGTTCTTAAATATCAGCAAGAG----- 594	QY 1421 ATTATATATGGCTTCCAAAGGGAGCAGGCAAGCGAGCTTTAAGAAACAACAACCCCTTCAAGT 1480	
QY	605 CATAAGCCCCCAAAAGTACGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACAGC 664	Db 1385 ACATTTCAACCTTAAAGAACCCAGCAGGCGAGGATTAAGAAACAACAACCCCTTCAAGT 1444	
Db	595 -----CAAGAGGTCATC 607	QY 1481 TCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGGGCTGTGGCTTA 1523	
QY	665 ACAGCCGACAGAACGAGCAGGACCAAGAAAGAAACGAGGTGGAAACATCTTCAGCG 724	Db 1445 TCTGTGTTCCACCTCAGGAGTCTCAGAAGAGAGCTGTGGCTTA 1487	
Db	608 AAAGCCAGAAAGGAAGCATCAGCAAGAAAGAAACGAGGAGGAGCATATTGAGTG 667		
QY	725 GCTTCAGCGCGAGTCTCTGGAACAGCCTTCCAGGTTGACGACAGACAGATAGTCAAA 784		
Db	668 GCTTCACCTTGAATTTCTTGGAAACATGCAATTCAGCGT---GGACAAAGCAGATAGCGAAA 724		
QY	785 ACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGGAGGCC 844		
Db	725 ACCTAAGAGAGAACGAGGAGGAGAACAGAGGGAGCCATTGTGACAGTGAAGAGGAGTC 784		
QY	845 TCAGAAATCTTTAGGCCCA-----GATAGAAAGAGACGCTGCCGACGAAGAGAGGAAT 895		

RESULT 13

ADH89244	ADH89244 standard; DNA; 1488 BP.
ID	ADH89244
XX	AC
XX	ADH89244;
XX	AC
DT	06-MAY-2004 (first entry)
XX	DE
XX	DE
XX	G. max glycinin A-la-B-x subunit DNA.
KW	de; gene; double stranded RNA; storage protein; 2S-albumen; 7S-globulin;
KW	11S/12S-globulin; zein-prolamine; homogenitistate metabolic pathway;
KW	pharmaceutical; plant; abiotic stress; fatty acid composition;
KW	lipid composition; oil composition; carbohydrate composition; colour;
KW	pigmentation; pathogen resistance; fruit ripening delay; aging;
KW	male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;
XX	caffeine; theophylline; threonine biosynthesis; glycine.
OS	Glycine max.
XX	WO2003078629-A1.
PN	25-SEP-2003.
XX	17-MAR-2003; 2003WO-BP002735.

XX	20-MAR-2002; 2002DE-01012892.	Db	251	ACACCAACCGGTCCCGAGGAAATCTATCAACCAAGAGGTAAAGGTATTTTGGCATGATAT	310
XX	(BADI) BASF PLANT SCI GMBH.	QY	245	TCCTGGTTGTCTAGACACTATGAAGAGCCTCACACAAAGGTGCTCGATCTCAGTCCC	304
XX		Db	311	ACCGGGTTGTCTAGACATTTGAAGACCTCA	344
XX	Kock M, Bauer J;	QY	305	AAAGACCACCAAGACGTCTTCAAGGAGAAAGCAAGCCAAACAGCAACGAGATAGTCACC	364
XX	WPI; 2003-803889/75.	Db	345	-----ACAACTCAACAAAGAGGACAAAGCAGCAGACCCACCAAGACCGTCACC	391
XX	P-PSDB; ADH89245.	QY	365	AGAGGTGACACCGTTTTCGATGAGGGTGATCTCATTTGACGTTCCCAACCGGTGCTTTCT	424
XX	Reducing expression of at least two target genes, useful e.g. for	Db	392	AGAAATCTATAACTTCAGAGAGGGTGATTTGATCGCAGTGGCTACTGCTGTTGCACTGT	451
XX	producing transgenic plants, using partly double-stranded interfering	QY	425	GGCTCTACAACGACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACAACA	484
XX	RNA.	Db	452	GGATGTACAACCAATGAAGACACTCTCTGTTGTTGCGGTTTCTATTTATTTGACCAACAGCT	511
XX	Claim 7; SEQ ID NO 19; 228pp; German.	QY	485	ACGACCAACAGCTTGTATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAACACGGAGCAAG	544
XX	This invention describes a novel method for reducing the expression of at	Db	512	TGGAGAACAGCTCGACAGATGCTCTAGGAGATTCTATCTTCTGCTGGAAACCAAGAGCAAG	571
XX	least two different endogenous target genes in a eukaryotic cell or	QY	545	AGTTCTTAAGTACACGACCAACAAAGCAGACAAAGCAGACAAAGAGCTTACCATATAGCC	604
XX	organism by introducing an RNA molecule that is at least partly double	Db	572	AGTTTCTAAATATATCAGCAAGAG-----	594
XX	stranded. The transcribed RNAs from at least two target genes have	QY	605	CATACAGCCGCAAAAGTCAAGCTCAGCTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGACAGC	664
XX	homology below 90% and the RNA molecule is formed as a single, self-	Db	595	-----	607
XX	complementary molecule. At least one of the double-stranded structures	QY	665	ACAGCCGACAGAAACGAGCAGGACGAAGAAGAAAGAAAGAGGTGGAAACATCTTTCAGCG	724
XX	formed from individual sense sequences has an even number of repeats of	Db	608	AAAGCCAGAAAGCAAGATCATGCAAGAAAGAAAGAAAGCAAGAGGAGCAGCATATTGAGTG	667
XX	21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At	QY	725	GCTTACGCGGAGTTCCTGGAACAAAGCTTCAGGTTGACGACAGACAGATAGTCAAA	784
XX	least two target genes are selected from different classes of storage	Db	668	GCTTACCTCCGAAATCTTTGGAACATGCAATTCAGCGT---GGACCAAGCAGATAGCGAANA	724
XX	protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-	QY	785	ACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGGAGGCC	844
XX	prolamine and at least one of the sense sequences is identical to storage	Db	725	ACCTACAGAGAGAAAGAAAGGGAAGAAACAGGAGCCATTGTGACAGTGAAGAGAGGTC	784
XX	protein sequences or genes in the homogenizate metabolic pathway or	QY	845	TCAGATCTTTAGCCCCA-----GATPAGAAAGAGACGTCCCGACGAAAGAGAGAAAT	895
XX	enzymes types, e.g. acetyl transacylases, chioesterases, (de)branching	Db	785	TCAGCGTGATAAAACCCACCACCGAGCAGCAGCAACAAAGACCCACGAGGAAGAGAAG	844
XX	expression systems, vectors and transgenic organisms are used for	QY	896	AGATGAAGATGAATA-----TGAATACGATGAGAGGATAGAGGCGTG	940
XX	preparation of pharmaceuticals, in biotechnological processes and plant	Db	845	AAAGAGAGGATGAGAAGCCACAGTCGCAAGGGTAAAGACAAACACTGCCAACCCCGCAG	904
XX	biotechnology, specifically in plants to improve protection against	QY	941	GCAGGGAAGCAGAGCAGCGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTA	1000
XX	abiotic stress, to modify composition and/or content of fatty acids,	Db	905	GAAAGCAAGCAAGCAGAGAAATGGCATTTGACGAGACCATATGCAACCATGAGACTTC	964
XX	lipids and oils, to modify carbohydrate composition, to alter colour or	QY	1001	AAAGAAACATTGGTAGAAACAGATCCCTTGACATCTTACAAACCTCAAGTGTGTTCACTCA	1060
XX	pigmentation, to reduce content of storage proteins, to increase	Db	965	GCACAAACATTGGCCAGACTTCATCACCTGACATCTTACAAACCTCAAGCGGTAGCGTCA	1024
XX	resistance to pathogens, to inhibit stem break, to delay fruit ripening	QY	1061	AAACTGCCAACATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTGTGATATG	1120
XX	or aging, to induce male sterility, to reduce content of toxic or	Db	1025	CAACGCCACCCAGCCTTGAATCTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTTG	1084
XX	unwanted components, to modify lignification and/or lignin content, to	QY	1121	GAAATCTTACAGGAATGCAATTTGTCGCTCACTTACAAACCAACGACACAGCATCA	1180
XX	susceptibility to shock, to increase synthesis of Vitamin E, to reduce	Db	1085	GATCTCTCGCAAGAAATGCAATGTTTCGTGCCACACTTACAACTGAACGCAACAGCATAA	1144
XX	contents of nicotine, caffeine or theophylline and to increase methionine	QY	1181	TATATCGATTGAGGGAACGGGCTCAGCTGCAAGTGTGTGACAGCAACGGCAACAGAGTGT	1240
XX	content, by reducing threonine biosynthesis. The method provides a rapid	Db	1145	TATACGCAATTGAATGGACGGGCAATTCATACAAAGTGTGAAATTCACACGGTGAGAGAGTGT	1204
XX	and efficient way of reducing gene expression, can inhibit more than one	QY	1241	ACGACGAGAGCTTCAAGAGGGTCAAGTGTGTCGACAGACTTTCGCGTGCCTG	1300
XX	target gene, prevents development of multiple phenotypes (since the	Db	1205	TTGATGGAGAGCTGCAAGAGGGACGGGTGCTGATCGTGCACAAAACCTTTGTTGGTGGCTG	1264
XX	transcription rate is the same for all RNA sequences, significantly				
XX	reducing the selection process required to produce an organism with				
XX	effective suppression of all target genes), avoids problems of epigenic				
XX	gene silencing, does not require synthesis of individual RNA sequences				
XX	and the method can be applied to plants with complex (polyploid) genomes.				
XX	No interference between the individual RNA sequences occur. This sequence				
XX	represents a target gene used in the method of the invention.				
XX					
XX	Sequence 1488 BP; 466 A; 368 C; 362 G; 292 T; 0 U; 0 Other;				
XX					
XX	Query Match 33.7%; Score 513; DB 10; Length 1488;				
XX	Best Local Similarity 62.7%; pred. No. 2.6e-136;				
XX	Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;				
QY	5 AGCAACGGAGGAGACGGCTGCGAGTTCGAGCGCTCAATGCGCAGAGACCTTGACAATC	64			
Db	71 AGCAGCCTCAGCAAAACGAGTGCAGATCCAAAAAATCTCAATGCCCTCAAAACCGGATAACC	130			
QY	65 GCATTGAATCAGAGGGGGTTACATTGAGACTTGGAAACCCCAACACCGAGGTTTCAAT	124			
Db	131 GTATAGAGTCAGAGAGGAGGCTCATTTGAGACATGGAACCCCTAACACAGCCATTCAAT	190			
QY	125 GCACCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTTCGTAGCCCTTTCT	184			
Db	191 GTCCCGGTGTTCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTTCGTAGACCTTCTCT	250			
QY	185 ACTCCAAATGCTCCCGAGGAGATCTTCATTCAGCAGGAAGGGGATACTTTGGTGTGATAT	244			

Db 659 AAAGCCAGAAAGGAAGCATCAGCAAGAAAGAAACGAAGGAGCGCATATTGAGTG 718
Qy 725 GCTTCAGCGCGGAGTTCTCGAAACAGCCTTCCAGGTTGACGACAGACAGATAGTGCAAA 784
Db 719 GCTTCACCTTGAATCTTGGAAACATGCTTACGCGT---GGACAAGCAGATAGCGAAAA 775
Qy 785 ACCTAAGAGCGGACGACGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGGGAGGCC 844
Db 776 ACCTACAAGAGAGAGAAAGAGGAGGAGACAAAGGAGGCCATTGTGACAGTGAAGGAGGTC 835
Qy 845 TCAGAACTCTTGAGCCCA-----GATAGAAAGAGACGTCGCGACGAAGAAGAGGAAT 895
Db 836 TGAGCGTGATAAAACCAACCCACGGACGACGAGCAAAAGACCCGAGAGAGAGAAAGAG 895
Qy 896 ACGATGAAGATGAATA-----TGAATACGATGAAGAGGATAGAAGGCGTG 940
Db 896 AAGAAGAGGATGAGAACCCACAGTGCAAGGGTTAAAGACAAACACTGCCAACGCCCCGAG 955
Qy 941 GCAGGGGAGCAGAGGCGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTA 1000
Db 956 GAAAGCCAAAGCAAAGCAGAAAGATGGCAATTGACGAGACCATATGCACCATGAGACTTC 1015
Qy 1001 AAAAGAACATTGTAGAAACAGATCCCTCGACATCTACAACTCAAGCTGTTCACTCA 1060
Db 1016 GCCACACATTTGGCCAGACTTTCATCCTGACATCTACAACTCAAGCGGTAGCGTCA 1075
Qy 1061 AAACTGCAACGATCTCAACCTTCTAATACTTAGTGGCTTGGACCTTAGTGTGAATATG 1120
Db 1076 CAACCGCCACAGCCTTGACTTCCAGCCTCTCGTGGCTCAGACTCAGTGTGCTGAGTTG 1135
Qy 1121 GAAATCTCTACAGAAATGCAATTTGTTGCTCCTCACTAAACCAACGACACAGCATCA 1180
Db 1136 GATCTCTCCGCAAGAAATGCAATGTTGTCGCACACTCAACCTGAACGCGAAACAGCATAA 1195
Qy 1181 TATATCGATTGAGGGACGGGCTCAGTGCAGTCTGGACAGCAACGCAACAGAGTGT 1240
Db 1196 TATACGCAATTGAATGGACGGGCATTGATACAGTGGTGAATTGCAACGGTGAGAGAGTGT 1255
Qy 1241 ACACGAGGAGCTTCAAGAGGGTCAAGTGTGTTGTTGGTGCACAGAACTTGCCTGCGCTG 1300
Db 1256 TTGATGGAGAGCTGCAAGAGGGACGGTGTGATCGTGCACAAACTTTGTGTGCTG 1315
Qy 1301 GAAAGTCCACAGAGCGAAGCTTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCA 1360
Db 1316 CAAAGTCAACAGAGTGACAACTTCGAGTATGTGTCAATTCAGACCAATGATACACCATGA 1375
Qy 1361 TAGCCAACTCGCGGTGAATACTCCGTCAATAGATAACCTGCCGAGGAGTGTGTTGCAA 1420
Db 1376 TCGGCACTTTGAGGGGCAAACTCATTTGTAACGCAATTACAGAGGAAGTATTCAGC 1435
Qy 1421 ATTATATGGCCTTCAAAGGGAGCAGCAAGGAGCTTAAGAAACAACACCCCTTCAAGT 1480
Db 1436 ACACCTTCAACCTAAAAGCCAGCAGCGCCAGGAGATAAAGAAACAACACCCCTTCAAGT 1495
Qy 1481 TCTTGGTTCCACGCTCTCAGCAGTCTCCGAGGGCTGTGGCTTA 1523
Db 1496 TCCTGGTTCCACCTCAGGAGTCTCAGAAAGAGAGCTGTGGCTTA 1538

Search completed: August 23, 2005, 21:26:22
Job time : 583.653 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 3574.43 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcacgaaccgagagagaa.....ctccgaggctgtggtcttaa 1524

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651.6	42.8	770	6	CD038329
2	604.4	39.7	725	6	CD038327
3	562.2	36.9	676	6	CD038254
4	554.6	36.4	710	6	CD038107
5	548.5	36.0	718	6	CD038332
6	487	32.0	553	7	CD087496
7	436	28.6	602	6	CD038301
8	417.6	27.4	532	6	CD038222
9	386.2	25.3	720	6	CD038043
10	373.6	24.5	452	7	CD087497
11	372	24.4	679	6	CD038144
12	330.4	21.7	906	5	BQ123409
13	321.6	21.1	974	5	BQ123283
14	316.2	20.7	886	6	CD038043
15	314	20.6	861	6	CD038043
16	310.6	20.4	810	6	CD038043
17	310.4	20.4	862	6	CD038043
18	310.2	20.4	907	5	BQ123241
19	307.8	20.2	822	5	BQ123014
20	305	20.0	787	6	CD038043
21	305	20.0	812	5	BQ123297
22	304.4	20.0	782	6	CD038043
23	304.4	20.0	879	5	BQ123329
24	303.6	19.9	824	6	CD038043

ALIGNMENTS

RESULT 1
CD038329
LOCUS
DEFINITION
CD038329 B05 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPP1007_B05 5', mRNA sequence.
ACCESSION
CD038329
VERSION
CD038329.1
KEYWORDS
EST.
SOURCE
Arachis hypogaea (peanut)
Arachis hypogaea
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 770)
Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G. and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..770
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPP1007_B05"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCVILX44). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent and ultrapurified (GIBCOBRL). mRNA was extracted and purified from

25 302.6 19.9 881 6 CA858490
26 298.4 19.6 790 6 CA858419
27 296.4 19.4 836 6 CA858453
28 296.2 19.4 812 4 B1311450
29 294 19.3 805 6 CA857850
30 293.4 19.3 827 6 CA858430
31 292.6 19.2 879 6 CA858252
32 292.4 19.2 789 4 B1310542
33 291.6 19.1 769 5 BQ124073
34 291 19.1 903 6 CA858778
35 290.8 19.1 879 6 CA858395
36 288.8 19.0 892 6 CA858816
37 288.6 18.9 746 4 B1311303
38 288.6 18.9 787 4 B1311483
39 288.2 18.9 793 6 CA857838
40 287.6 18.9 541 6 CD038610
41 287.4 18.9 752 6 CA858702
42 287.2 18.8 755 6 CA858724
43 286.4 18.8 756 6 CA857897
44 286.2 18.8 811 5 BQ123350
45 285.2 18.7 796 6 CA857752

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match		42.8%;	Score 651.6;	DB 6;	Length 770;
Best Local Similarity		92.9%;	Pred. No. 1.7e-168;		
Matches 694;		Conservative 0;	Mismatches 44;	Indels 9;	Gaps 1;
QY	391	GATCTCATGTGCAGTTCCACCGGTGTGGCTTCTGCGCTCTCAACGACGACGACTGAT	450		
DB	33	GATCTCATNGNNNACCGGCTCGAGGAATTTCGGCACGAGGCAACGACCATGACACTGAT	92		
QY	451	GTGTGTTGCTGTTCTCTTACTGACACCAACACGACGACGAGCTTGATCAGTTCCCC	510		
DB	93	GTGTGTTGCTGTTCTCTTACTGACACCAACACGACGACGAGCTTGATCAGTTCCCC	152		
QY	511	AGGAGATTCAATTTGGCTGGGAACACGGGACGAGAGTTCCTTAAGGTACCAGCAACAAAGC	570		
DB	153	AGGAGATTCAATTTGGCTGGGAACACGGGACGAGAGTTCCTTAAGGTACCAGCA	205		
QY	571	AGCAAGACGACGACGAGAGCTTACATATAGCCCATAGCCCGCAAGTCTAGCTCTAGA	630		
DB	206	--ACAAAGCAGACGACGAGAGCTTACATATAGCCCATAGCCCGCAAGTCTAGCTCTAGA	263		
QY	631	CAAGAAGAGGCTCAATTTAGCTTCAGGACACGACGCGGACGAGCAACGAGCAGGACAA	690		
DB	264	CAAGAAGAGGCTCAATTTAGCTTCAGGACACGACGCGGACGAGCAACGAGCAGGACAA	323		
QY	691	GAAGAAGAAAACGAAGGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCCTGGAAACAA	750		
DB	324	GAAGAAGAAAACGAAGGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCCTGGCAACAA	383		
QY	751	GCCTTCAGGTTGACACACACAGATAGTGCAAAACCTAAGAGGCGAGACCGAGGTGAA	810		
DB	384	GCCTTCAGGTTGACACACACAGATAGTGCAAAATCTAAGAGGCGAGAACGAGAGTGAG	443		
QY	811	GAAGAGGAGCCATTGTGACAGTGAGGAGGCGCTCAGAAATCTTGAAGCCACAGATAGAAAG	870		
DB	444	GAAGAGGAGCCATTGTGACAGTGAGGAGGCGCTCAGAAATCTTGAAGCCACAGATAGAAAG	503		
QY	871	AGACGTGCCGACGAAGAAGGGAATACGATGAAGATGAATATGAATACGATGAAGAGGAT	930		
DB	504	AGAGGTGCCGACGAAGAAGGGAATACGATGAAGATGAATATGAATACGATGAAGAGGAT	563		
QY	931	AGNAGCGTGGCAGGGAAGCAGAGCGGCGGAATGGTATTGAAGACAGATCTGCACC	990		
DB	564	AGNAGCGTGGCAGGGAAGCAGAGCGGCGGAATGGTATTGAAGACAGATCTGCACA	623		
QY	991	GCAAGTGCTAAAAGAACATTGGTAGAAAACAGATCCCTGACATCTTACAAACCTCAAGCT	1050		
DB	624	GCNAGTGTAAAAGAACATTGGTAGAAAACAGATCCCTGACATCTTACAACTCTCAAGCT	683		
QY	1051	GGTTCTACTCAAAACCTGCCACGATCTCAACCTTTCTAATATCTTAGTGGCTTGACCTTACT	1110		
DB	684	GGTTCTACTCAAAACCTGCCACGATCTCCACCTTCTAATCTTAGTGGCTTGACCTTACT	743		
QY	1111	GCTGAATATGGAATCTCTACAGGAAT	1137		
DB	744	GCTGAATATGGAATCTCTACAGGNAT	770		

RESULT 2

CD038327

LOCUS

CD038327 725 bp mRNA linear EST 07-MAY-2003

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

UTPPI008_F09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI008_F09 5', mRNA sequence.
CD038327
CD038327.1 GI:30420165
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 725)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

REFERENCE

AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES

source

1..725
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI008_F09"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="xL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 39.7%; Score 604.4; DB 6; Length 725;
Best Local Similarity 92.8%; Pred. No. 1.9e-155;
Matches 648; Conservative 0; Mismatches 41; Indels 9; Gaps 1;

QY	2	GGCAGCAACCGGAGGAGAACGCGTCCAGCTTCCAGCGCTCAATCGCAGACCTTGACA	61
DB	37	GGCAGCAGCGGAGGAGAAATGCGTCCAGTTCAGCGCTCAATCGCAAGGCTTGACA	96
QY	62	ATCGCATTAATCAGAGGCGGTACATTGACATTCGAAACCCCAACACGAGGATTCG	121
DB	97	ACCGCATTTGAATCGGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGATTCG	156
QY	122	AATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCCTT	181
DB	157	AATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCCTT	216

```

QY 182 TCTACTCAATGCTCCCGAGGATCTTCATCCAGCAAGGAAGGGATACCTTTGGTTGA 241
Db 217 TCTACTCAATGCTCCCGAGGATCTTCATCCAGCAAGGAAGGGATACCTTTGGTTGA 276
QY 242 TATTCCCTGGTGTCTTAGACACTATGAAGAGCTTCACACAAAGGTGCTGCATCTCAGT 301
Db 277 TATTCCCTGGTGTCTTAGACACTATGAAGAGCTGTCACAAAGGAGCGCGACATCAGT 336
QY 302 CCCAAGACCAACCAAGAGCTTCCAGGAGAGCAAGCCAAAGCCAAACAGCAAGATAGTC 361
Db 337 CCCAAGACCAACCAAGAGCGTTTTTCAAGGACAAGCAAGCCAAAGCCAAACAGCAAGATAGTC 396
QY 362 ACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTTCCACCGGTGTGCTT 421
Db 397 ACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTTCCACCGGTGTGCTT 456
QY 422 TCTGGCTCTCAACGACCAAGCAGACACTGATGTGTTGCTGTTCTTCTTACTGACACCAACA 481
Db 457 TCTGGATGTACAAGCAACCATGACACTGATGTGTTGCTGTTCTTCTTACTGACACCAACA 516
QY 482 ACAACGACAACAGCTTGATCAGTTTCCCGAGGAGATTCAATTTGGCTGGGAACACGGAGC 541
Db 517 ACAACGACAACAGCTTGATCAGTTTCCCGAGGAGATTCAATTTGGCTGGGAACACCGAGC 576
QY 542 AAGAGTTCTTAAGGTACCAACCAACAGCAGACAAAGCAGCAAGAGCTTACCATATA 601
Db 577 AAGAGTTCTTAAGATACCAAGCA-----ACAAAGCAGCAAGAGCTTACCATATA 627
QY 602 GCCATACAGCCCGCAAGTCAGCTTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGAC 661
Db 628 GCCATACAGCCCGCAAGTCAGCTTAACCAAGAGACCGTGAATTTAGCCCTCGAGGAC 687
QY 662 AGCAGCCCGCAGAGAACGAGCAGGACCAAGAAAGAA 699
Db 688 AGCAGCCCGCAGAGAACGAGCAGGACCAAGAAAGAA 725

RESULT 3
CD038254 LOCUS
DEFINITION
  CD038254 676 bp mRNA linear EST 07-MAY-2003
  Arachis hypogaea cDNA clone UTPPI004_F06 5', mRNA sequence.
ACCESSION
  CD038254
VERSION
  CD038254.1 GI:30420092
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
REFERENCE
  1 (bases 1 to 676)
  Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
  and Lynch, R. E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3.
FEATURES
    source
      Location/Qualifiers
        1..676
        /organism="Arachis hypogaea"
        /mol_type="mRNA"
        /cultivar="A13"
        /db_xref="taxon:3818"
        /clone="UTPPI004_F06"
        /tissue_type="Immature pods"
        /dev_stage="R6"

```

```

/lab_host="Xli-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV11XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."
ORIGIN
Query Match 36.9%; Score 562.2; DB 6; Length 676;
Best Local Similarity 92.2%; Pred. No. 8.4e-144;
Matches 630; Conservative 0; Mismatches 38; Indels 15; Gaps 3;
QY 538 GAGCAAGAGTTCTTAAGGTACCAAGCAAAAGCAGACAAAGCAGAGCAAGAGCTTACCA 597
Db 3 GAGCAAGAGTTCTTAAGATACCAAGCAAAAGCAGACAAAGCAGAGCAAGAGCTTACCA 62
QY 598 TATAGCCCATACAGCCCGCAAAAGTCAAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGA 657
Db 63 TTAAGCCCATACAGCCC-----TCAGCCTGGACAAGAGACCGTGAATTTAGCCCTCAA 116
QY 658 GGACAGCAGCGCGCAGAGAACGAGCAGCAGACAGAGAGAGAAACAGAGGTGGAACATC 717
Db 117 GGACAGCAGCGCGCAGAGAACGAGCAGCAGACAGAGAGAAACAGAGGTGGAACATC 176
QY 718 TTCAGCGGCTTCACCGCGGAGTTCTTCGGAACAGCCCTTCAGGTTGACGACAGACAGATA 777
Db 177 TTCAGCGGCTTCACGTCGGAGTTCTTCGCAAGCCCTTCAGGTTGACGACAGACAGATA 236
QY 778 GTGCAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGAGCCCATTTGTGACAGTAGG 837
Db 237 GTGCAAAATCTTAAGAGCGGAGAACGAGAGTGAAGAGACAGGAGCCCATTTGTGACAGTAGG 296
QY 838 GGAGGCGCTCAGATCTTGAGCCGACAGATAGAAAGAGAGCGTCCGACGAAAGAGGAGATAC 897
Db 297 GGAGGCGCTCAGATCTTGAGCCGACAGATAGAAAGA---GTCCCCGACGAAAGAGGAGATAC 353
QY 898 GATGAAGATGAATATGAATACGATG-----AAGAGGATAGAGGCGTGGCAGGGGAAGC 951
Db 354 GATGAAGACGATATGCTGAGAGGAGAGGCAACAGATAGAGGCGTGGCAGGGGAGC 413
QY 952 AGAGGCGAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGACATT 1011
Db 414 AGAGGCGAGGGCAATGGCATTTGAGGAGACCATCTGCACCGCAACTGTTAAAAAGACATT 473
QY 1012 GGTAGAAACAGATCCCTCGACATCTCAACCCCTCAAGCTGGTTCACTCAAAATGCCAAC 1071
Db 474 GGTAGAAACAGATCCCTCGACATCTCAACCCCTCAAGCTGGTTCACTCAAAATGCCAAC 533
QY 1072 GATCTCAACCTTCTAATACTTAGTGGCTTGGACCTTAGTGTCTGAATATGGAATCTCTAC 1131
Db 534 GAGCTCAACCTTCTAATCTTAGTGGCTTGGACTTAGTGTGAATATGGAATCTCTAC 593
QY 1132 AGGAATGCATTTGTTGCTCGCTCACTACAAACCAACGCAACAGCAGCATCATATCGAATTG 1191
Db 594 AGGAATGCATTTGTTGCTCGCTCACTACAAACCAACGCAACAGCAGCATCATATCGAATTG 653
QY 1192 AGGGGACGGGCTCACGTCGAAGT 1214

```

```

Db      654 AGGGGACGGGCTCATGTGCAAGT 676
|||||
CD038107      710 bp mRNA linear EST 07-MAY-2003
UTPPI002_E08 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI002_E08 5', mRNA sequence.
CD038107
CD038107.1 GI:30419945
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 710)
Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd, Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..710
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI002_E08"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV1XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

RESULT 4
LOCUS      CD038107
DEFINITION UTPPI002_E08 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION  CD038107
VERSION     CD038107.1
KEYWORDS    EST.
SOURCE      Arachis hypogaea (peanut)
ORGANISM    Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 710)
Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd, Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..710
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI002_E08"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV1XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match      36.4%; Score 554.6; DB 6; Length 710;
Best Local Similarity 98.3%; Pred. No. 1.1e-141;
Matches 571; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY      945 GGGAGCAGAGCGGCGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAA 1004
|||||
Db      1 GGGAGCAGAGCGGCGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAA 60
|||||
QY      1005 GAACATTGGTAGAACAGATCCCTCGACATCTACAACCCCTCAAGCTGTCTCACTCAAAAC 1064
|||||

```

/clone.lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapur (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA GigaPack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using GigaPack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 36.0%; Score 548.6; DB 6; Length 718;
Best Local Similarity 93.4%; Pred. No. 4.9e-140;
Matches 621; Conservative 0; Mismatches 29; Indels 15; Gaps 4;
QY 875 GTCCCGACGAAGAGGATACGATGAAGATGAATGAATACGATGAAG----- 927
DB 1 GTCCCGACGAAGAGGATACGATGAAGATGAATGAATGAATGAAGAGGAGGC 60
QY 928 -----GATAGAGCGTGGCAGGGGAGCAGAGCGGGGATGGTATTGAAGACGA 982
DB 61 AACAAAGATAGAGCGTGGCAGGGGAGCAGAGCGGCAATGGCATTTGAGGACCA 120
QY 983 TCTGCACCGCAAGTCTTAAAGAAACATTCGTAGAACACATCCCTGACATCTACAACC 1042
DB 121 TCTGCACCGCAAGTCTTAAAGAAACATTCGTAGAACACATCCCTGACATCTACAACC 180
QY 1043 CTCAGCTGGTTCACTCAAAACCTGCCAACGATCTCAACCTTCTAATACTTAGTGGCTTG 1102
DB 181 CTCAGCTGGTTCACTCAAAACCTGCCAACGATCTCAACCTTCTAATACTTAGTGGCTTG 240
QY 1103 GACCTAGTCTGAATATGGAATCTCTACAGGAATGATTTTGTGCTCACTACACA 1162
DB 241 GACTTAGTCTGAATATGGAATCTCTACAGGAATGATTTTGTGCTCACTACACA 300
QY 1163 CCAACGCACACATCATATATCGATTGAGGGGACGGGCTCACGTGCAAGTGTGGACA 1222
DB 301 CGAACGCACACATCATATATGCAATTGAGGGGACGGGCTCACGTGCAAGTGTGGACA 360
QY 1223 GCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTGTGGTGGCCAC 1282
DB 361 GCAACGGCAACAGAGTGTGTGACGAGGAGCTTCAAGAGGGTCAAGTGTGGTGGCCAC 420
QY 1283 AG-AACTTCCCGTGGTGAAGTCCAGAGCGAGAACTTCAATACGTGGCAATCAAG 1341
DB 421 AGAACTTCCCGTGGTGGTGAAGTCCAGAGCGAGAACTTGAATACGTGGCAATCAAG 480
QY 1342 ACAGACTCAAGGGCCAGCATAGCAACCTCCGCGGTGAAACTCCCGTCATAGATAACCTG 1401
DB 481 ACAGACTCAAGGGCCAGCATAGCAACCTAGCCGGTGAAGAACTCTCTTCATAGATACTTG 540
QY 1402 CC-GGAGGAGGTGGTGAAGATTCATATGGCTCCAAAGGGAGCAGGCAAGGAGCTTAA 1460
DB 541 CCGGAGGAGGTGGTGAAGATTCATATGGCTCCAAAGGGAGCAGGCAAGGAGCTTAA 600
QY 1461 GAACAACAA-CCCTTCAAGTCTTGTTCACCGTCTCAGAGTCTCCGAGGGCTGTGG 1519
DB 601 GAACAACAAACCCCTTCAAGTCTTGTTCACCGTCTCAGAGTCTCTGAGGGCTGTGG 660
QY 1520 CTTAA 1524
|||||

Db 661 CTTAA 665

RESULT 6
CO897496LOCUS
DEFINITION

CO897496 553 bp mRNA linear EST 13-AUG-2004
EST00001 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.

ACCESSION
VERSION

CO897496
CO897496.1 GI:512337286

KEYWORDS

EST.

SOURCE

Arachis hypogaea (peanut)

ORGANISM

Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 553)

AUTHORS

Yan, Y.S., Wang, L. and Huang, S.Z.

TITLE

Expressed sequence tags of mid-matured cotyledons of peanut

JOURNAL

Unpublished (2004)

COMMENT

Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (860) 02084036592
Email: YYS9803@yahoo.com.cn

Seq primer: ptriplex2 Forward.

FEATURES

Location/Qualifiers

1..553

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="Yueyou 523"

/db_xref="taxon:3818"

/tissue_type="Cotyledons"

/dev_stage="Mid-matured stage"

/lab_host="E.coli BM25.8"

/clone_lib="Peanut Lambda Express library"

/note="Organ: Seed; Vector: lambdaTriplex2"

ORIGIN

Query Match 32.0%; Score 487; DB 7; Length 553;
Best Local Similarity 93.6%; Pred. No. 4.8e-123;
Matches 508; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 39 CCTCAATGCGCAGAGACCTGACAAATCGCATTCGAATCAGAGGGCGTTACATTGAGACTTG 98
DB 1 CCTCAATGCGCAAGGCGCTGACAAACCGCATTCGAATCGAGGGCGTTACATTGAGACTTG 60
QY 99 GAACCCCAACACAGAGGTTGAAATGCGCGCGCGCTCTCTCGCTTAGTCTCTCCG 158
DB 61 GAACCCCAACACAGAGGTTGAAATGCGCGCGCGCTCTCTCGCTTAGTCTCTCCG 120
QY 159 CGCAACGCGCTTCTGAGGCTTTCTACTCCAAATGCTCCCGAGGAGATCTTCAAGCA 218
DB 121 CGCAACGCGCTTCTGAGGCTTTCTACTCCAAATGCTCCCGAGGAGATCTTCAAGCA 180
QY 219 AGGAAGGGGATATCTTTGGGTTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCTCA 278
DB 181 AGGAAGGGGATATCTTTGGGTTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCTCG 240
QY 279 CACACAGGTTGTCGATCTCAGTCCCAAGAGCACCAGACGCTCTCAAGGAGAGACCA 338
DB 241 ACAACAGAGCGCCGACATCAGTCCCAAGAGCACCAGACGCTTTCAAGGAGAGACCA 300
QY 339 AAGCCAAACAGCAACAGATAGTACCAGAGGTTGACCGTTTCGATGAGGTTGATCTCAT 398
DB 301 AAGCCAAACAGCAACAGATAGTACCAGAGGTTGACCGTTTCGATGAGGTTGATCTCAT 360
QY 399 TGCAGTTCCACCGGTTGTTGCTTTCTGGCTCTACACAGCAGCAGCACTGATGTTGTCG 458
DB 361 TGCAGTTCCACCGGTTGTTGCTTTCTGGATGTACACAGCAGCAGCACTGATGTTGTCG 420

```

QY 459 TGTTCCTCTTACTGACACCAACAAACGACCAACCGAGCTTGATCAGTTCCCCAGGAGATT 518
Db 421 TGTTCCTCTTACTGACACCAACAAACGACCAACCGAGCTTGATCAGTTCCCCAGGAGATT 480

QY 519 CAATTTGGCTGGGAACACGAGCAAGAGTCTTAAAGGTACCGACCAACAAGACGACAAAG 578
Db 481 CAATTTGGCTGGGAACACGAGCAAGAGTCTTAAAGATACCGACCAACAAGACGACGAAG 540

QY 579 CAG 581
Db 541 AAG 543

RESULT 7
CD038301
LOCUS
DEFINITION
  UTPII005_C08 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
  Arachis hypogaea cDNA clone UTPII005_C08 5', mRNA sequence.
ACCESSION
  CD038301
VERSION
  CD038301.1 GI:30420139
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
    Aeschynomeneae; Arachis.
REFERENCE
  1 (bases 1 to 602)
  Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
  and Lynch,R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  CONTACT: Baozhu Guo
  MOLECULAR GENETICS
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3
  Location/Qualifiers
    1..602
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPII005_C08"
      /tissue_type="Immature pods"
      /dev_stage="R6"
      /lab_host="XLI-blue"
      /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
      (UTPP)"
      /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
      cDNA library was constructed from peanut cultivar A13
      (NCV1XAR4). A13 has resistance to Aspergillus infection
      and drought tolerance. The immature pods that developed to
      R6 stage were collected from different plants, and placed
      into liquid N2 immediately and stored in -80oC freezer.
      Total RNA was isolated with TRIZOL-Reagent
      ultrapure (GIBCOBRL). mRNA was extracted and purified from
      total RNA (Promega). cDNA synthesis and library
      construction followed the protocol of by ZAP-cDNA Gigapack
      III Gold cloning kit (Stratagene). The cDNA above 500bp
      were collected after size-fraction. The inserts were
      directionally cloned into Uni-ZAP XR vector using XhoI
      EcoRI sites adapters. The lambda library was packed into
      phages using Gigapack III Gold (Stratagene). The
      un-amplified library was used to excise phaluescript
      phagemids from the Uni-ZAP XR vector, and the phagemids
      was used to transform the host bacteria SOLR. The library
      was constructed by Dr. Meng Luo and Dr. Phat Dang."

FEATURES
  source
    1..602
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPII005_C08"
      /tissue_type="Immature pods"
      /dev_stage="R6"

```

```

Query Match      28.6%; Score 436; DB 6; Length 602;
Best Local Similarity 95.7%; Pred. No. 6.2e-109;
Matches 448; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1057 CTCAAAACCTGCCAAGATCTCAACCTTCTAATACATTAGTGGCTTGGACCTAGTGTGAA 1116
Db 1 CTCAAAACCTGCCAAGAGCTCAACCTTCTAATCCTTAGGTGGCTTGGACTTAGTGTGAA 60

QY 1117 TATGGAATCTCTACAGGAATGCATTGTTTGGCTCACTACAACACCAACACACACAGC 1176
Db 61 TATGGAATCTCTACAGGAATGCATTGTTTGGCTCACTACAACACGACACACAGC 120

QY 1177 ATCATATATTCGATTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGA 1236
Db 121 ATCATATATTCGATTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGA 180

QY 1237 GTGTACGAGGAGGACTTCAAGAGGGTCAAGTGTGTGTGGTGGCCACAGAACTTTCGCGTTC 1296
Db 181 GTGTTCGACGAGGAGCTTCAAGAGGTCAAGAGGTCAAGTGTGTGTGGTGGCCACAGAACTTTCGCGTTC 240

QY 1297 GCTGGAAGTCTCCAGAGCGAGAACTTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCC 1356
Db 241 GCTGGAAGTCTCCAGAGCGAGAACTTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCC 300

QY 1357 AGCATAGCCAACTCGCGGTGAAACCTCCGTCATAGATAAATCTCCGAGAGAGTGGTT 1416
Db 301 AGCATAGCCAACTAGCGGTGAAACCTCCGTCATAGATAAATCTCCGAGAGAGTGGTT 360

QY 1417 GAAATTCATATGGCTTCCAAAGGAGGAGGCAAGGAGCTTAAAGACACAAACCCCTTC 1476
Db 361 GAAATTCATATGGCTTCCAAAGGAGGAGGCAAGGAGCTTAAAGACACAAACCCCTTC 420

QY 1477 AAGTTCCTTCTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 421 AAGTTCCTTCTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 468

RESULT 8
CD038222
LOCUS
DEFINITION
  UTPII004_B11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
  Arachis hypogaea cDNA clone UTPII004_B11 5', mRNA sequence.
ACCESSION
  CD038222
VERSION
  CD038222.1 GI:30420060
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
    Aeschynomeneae; Arachis.
REFERENCE
  1 (bases 1 to 532)
  Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
  and Lynch,R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  CONTACT: Baozhu Guo
  MOLECULAR GENETICS
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3
  Location/Qualifiers
    1..532
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPII004_B11"
      /tissue_type="Immature pods"
      /dev_stage="R6"

```

/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 27.4%; Score 417.6; DB 6; Length 532;
Best Local Similarity 96.8%; Pred. No. 7.3e-104; Indels 0; Gaps 0;
Matches 426; Conservative 0; Mismatches 14;

QY 1085 TAATACCTAGGTGGCTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCAATTGT 1144
DB 1 TAATCCTTAGGTGGCTGGACTAGTCTGAATATGGAATCTCTACAGGAATGCAATTGT 60
QY 1145 TTGTGCTCTACTCAACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC 1204
DB 61 TTGTGCTCTACTCAACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC 120
QY 1205 ACCTGCAAGTCGTGGACAGCAACGCAAGAGTGACGAGGAGGCTTCAAGAGGCTC 1264
DB 121 ACCTGCAAGTCGTGGACAGCAACGCAAGAGTGACGAGGAGGCTTCAAGAGGCTC 180
QY 1265 ACCTGCTTTGTGGCCACAGAACTTCGCCCTGCTGGAAAGTCCACAGAGGAGAACTTCG 1324
DB 181 ACCTGCTTTGTGGCCACAGAACTTCGCCCTGCTGGAAAGTCCACAGAGGAGAACTTCG 240
QY 1325 AATAGTGGCAATTCAGACAGACTCAAGGCCGACGATAGCAACCTTCGCCGCTGAAAACT 1384
DB 241 AATAGTGGCAATTCAGACAGACTCAAGGCCGACGATAGCAACCTTCGCCGCTGAAAACT 300
QY 1385 CCCTCATAGATACTGCCGAGGAGGTGCTGCAATTCATATGCTTCAAGAGGAGC 1444
DB 301 CCATCATAGATACTGCCGAGGAGGTGCTGCAATTCATATGCTTCAAGAGGAGC 360
QY 1445 AGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGT 1504
DB 361 AGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGT 420
QY 1505 CTCGAGGGCTGTGGCTTAA 1524
DB 421 CTCGAGGGCTGTGGCTTGA 440

RESULT 9
CD038043
LOCUS CD038043 720 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI001_F09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI001_F09 5', mRNA sequence.
ACCESSION CD038043
VERSION CD038043.1 GI:30419881
SOURCE EST.
ORGANISM Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

rosids; eurousids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 720)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES

source

Location/Qualifiers
1..720
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPP1001_F09"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 25.3%; Score 386.2; DB 6; Length 720;
Best Local Similarity 96.8%; Pred. No. 3.8e-95;
Matches 394; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1118 ATGGAATCTCTACAGGAATGCAATGTTTGTGCTCACTACACACCAACGACACAGCA 1177
DB 1 ATGGAATCTCTACAGGAATGCAATGTTTGTGCTCACTACACACGACGACACAGCA 60
QY 1178 TCATATATCGATTGAGGGACGGGCTCAGTGCAGTCTGTGGACACCAACGCAACAGAG 1237
DB 61 TCATATATCGATTGAGGGACGGGCTCAGTGCAGTCTGTGGACACCAACGCAACAGAG 120
QY 1238 TGTACGACGAGGAGCTTCAAGAGGGTCAGTGTGTTGTCGCCAAGAATTCGCCGTG 1297
DB 121 TGTACGACGAGGAGCTTCAAGAGGGTCAGTGTGTTGTCGCCAAGAATTCGCCGTG 180
QY 1298 CTGGAAGTCCACGACGAGAACTTCGAATACGTGCAATTCGAAGCAGACTCAAGGCCCA 1357
DB 181 CTGGAAGTCCACGACGAGAACTTCGAATACGTGCAATTCGAAGCAGACTCAAGGCCCA 240
QY 1358 GCATAGCCAACTCGCCGCTGAAATCTCCGTATAGATAACCTGCCGAGGAGGTGTTG 1417
DB 241 GCATAGCCAACTCGCCGCTGAAATCTCCGTATAGATAACCTGCCGAGGAGGTGTTG 300
QY 1418 CAATTCATATGCGCTCCAAAGGAGGACGCGCAAGGCTTAAGACACACACCCCTTCA 1477

301 CAAATTCATATGCTCCCAAGGAGCAGGCAAGGAGCTTAAAGAACAAACCCCTTCA 360
1478 AGTTCTTCGTTCCACCGTCTCACAGCTCCCGAGGCTGTGGCTTAA 1524
361 AGTTCTTCGTTCCACCTTTTTCAGAGTCTCCGAGGCTGTGGCTTAA 407

RESULT 10
CO897497

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CO897497 452 bp mRNA linear EST 13-AUG-2004
EST00002 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.
CO897497
CO897497.1 GI:51237287
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 452)
Yan, Y.S., Wang, L. and Huang, S.Z.
Expressed sequence tags of mid-matured cotyledons of peanut
Unpublished (2004)
Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (86) 02084036592
Email: yys9803@yahoo.com.cn
Seq primer: pTriplex2 Forward.
Location/Qualifiers
1. .452
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/issue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/notes="Organ: Seed; Vector: lambdaTriplex2"

FEATURES
source

ORIGIN

Query Match 24.5%; Score 373.6; DB 7; Length 452;
Best Local Similarity 91.7%; Pred. No. 1e-91;
Matches 407; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

2 GGCAGCAACCGGAGGAGAACGCGTGCAGTTCAGCGCCTCAATGCGCAGAGACCTTGACA 61
9 GGCAGCAGCGGAGGAGATGCGTGCAAGTTCAGGCGCTCAATGCGCAGAGCCTTGACA 68
62 ATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACACCCAGGAGTTCC 121
69 ACCGCTTGAATCGGAGGCGGTTACATTGAGACTTTGGAAACCCCAACACCCAGGAGTTCC 128
122 AATGCGCGCGCGTCTCTCGCTTAGTCTCTCGCGCAACGCGCTTCGTAGGCGCTT 181
129 AATGCGCGCGCGTCTCGCTTAGTCTCTCGCGCAACGCGCTTCGTAGGCGCTT 188
182 TCTACTCCAAATGCTCCCGAGGAGATCTTCAATCCAGCAAGGAGGATCTTTGGGTTGA 241
189 TCTACTCCAAATGCTCCCGAGGAGATCTTCAATCCAGCAAGGAGGATCTTTGGGTTGA 248
242 TATTCCCTGTTGCTCTAGACACTATGAGAGGCTTACACACAAAGTCTGTCGATCTCAGT 301
249 TATTCCCTGTTGCTCTAGACACTATGAGAGGCTTACACAAAGTCTGTCGCGCATCAGT 308
302 CCACAAAGACCCCAAGACGCTCTCCAGGAGAGAGACCAACCAACGCAACGCAACGCA 358
309 CGCAAGAGACCCCAAGACGCTTTTGAAGGAGAGAGACCAACCAAGCAACGCAACGCA 368

QY 359 GTCCACCAAGGTGCACCGCTTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTG 418
Db 369 GTCCACCAAGAGTGGCGCTTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTG 428

QY 419 CTTTCTGGCTTACCAACGACACG 442
Db 429 CTCTCTGGATGTACAACGACCATG 452

RESULT 11
CD038144
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CD038144 679 bp mRNA linear EST 07-MAY-2003
UTPPI003 All USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI003_All 5', mRNA sequence.
CD038144
CD038144.1 GI:30419982
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 679)
Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1. .679
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI003_All"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="Xll-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

FEATURES
source

ORIGIN

Query Match 24.4%; Score 372; DB 6; Length 679;
Best Local Similarity 92.9%; Pred. No. 3.2e-91;
Matches 390; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1105 CCTAGTCTGTAATGGAATCTCTACAGGAATGCATTTGTTGCGTCTACTACAACACC 1164
 Db |||||
 46 CCTGAGCTCGAGGAATTCGCGACGAGGGGAATGCATTTGTTGTCCTCTACTACAACAGC 105
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1224
 Db |||||
 106 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCATGTGCAAGTCGTGGACAGC 165
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1284
 Db |||||
 166 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 225
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1344
 Db |||||
 226 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 285
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1404
 Db |||||
 286 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 345
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1464
 Db |||||
 346 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 405
 Qy AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 1524
 Db |||||
 406 AAGCCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTGGACAGC 465

RESULT 12

BQ123409
 LOCUS BQ123409.1 GI:20175371
 DEFINITION EST608985 GLSD Medicago truncatula cDNA clone pGLSD-32121, mRNA sequence.

ACCESSION BQ123409
 VERSION BQ123409.1
 KEYWORDS Medicago truncatula (barrel medic)
 SOURCE Medicago truncatula
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 906)
 AUTHORS Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., and Fraser, C.M.
 TITLE ESTs from late stage developing seeds of Medicago truncatula
 JOURNAL Unpublished (2002)

COMMENT Contact: Grusak, M.A.
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA
 Tel: 713 798 7044
 Fax: 713 798 7078

Email: mgrusak@bcm.tmc.edu
 TIGR sequence name: MTRAQ59TK More information is available at:
 www.medicago.org

Seq primer: SKmd (CTA GAA CTA gtg gat CC).

FEATURES
 source

1..906
 /location/Qualifiers
 1..906
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pGLSD-32121"
 /tissue_type="Immature seeds"
 /dev_stage="25 to 35 days after pollination"
 /lab_host="XLOLR"
 /clone_lib="GLSD"
 /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from

greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

ORIGIN

Query Match 21.7%; Score 330.4; DB 5; Length 906;
 Best Local Similarity 63.5%; Pred. No. 1.1e-79;
 Matches 540; Conservative 0; Mismatches 301; Indels 9; Gaps 2;
 Qy 680 GAGCAGGACAAGAGAAAGAAACGAAAGGTGGAAACATCTTTACGCGGCTTCCACGCCGAGT 739
 Db |||||
 47 GAGGAAGGGAACAACAAGAAATATGAAGGAGGCAACATTTTCAGTGGCTTCAAGAGGCGATT 106
 Qy 740 TCCTGGACAAAGCCTTCCAGGTTGACGACAGACAGATAGTGCAAAACCTTAAGAGGCGAGA 799
 Db |||||
 107 TCTTGGAAAGATGCATTTGAACGTGA---ACAGGCATATAGTTGAAACATTTCAAGGCGAGGA 163
 Qy 800 CCGAGAGTGAAGAGAGGAGGAGCCATTTGTGACGTGAGGGGAGGCCTCAGAAATCTTTGAGCC 859
 Db |||||
 164 ATGAACACAGGAGAGAGGAGCCATTTGTCAAAGTGAAGAGGAGACTCAGCATCATGAGCC 223
 Qy 860 C-----AGTAGAAGAGAGACGTGCCGACGGAAGAGGAATACGATGAAGATGAATG 913
 Db |||||
 224 CTCGAGAGAGACAACAACGCCACCCAGCAGACAGATGAAGATGAAGATGAAGAGG 283
 Qy 914 AATAAGATGAAGAGATAGAAAGCGGTGGCAGGGGAAAGCAGAGCGGGGAAATGGTATTG 973
 Db |||||
 284 ATGAGTGGAGGCCACATCATCAGAAAGAGAGAGAAAGAGAAAGGCAACACGCTTTG 343
 Qy 974 AAGAGACGATCTGCACGCGCAAGTGTCTAAAGAACATTTGTTAGAAACAGATCCCTGCACA 1033
 Db |||||
 344 CGGAACAATTTGCACTGCCAGGCTTCCAGCAACATGTTCTCATTTTCATCACTGACA 403
 Qy 1034 TCTACAACCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAACCTTTCTAATCTTA 1093
 Db |||||
 404 TCTACAACCTCAAGCTGGTTAGAAATCAAAATATACACGATCTCGACCTCCAGCTCTCA 463
 Qy 1094 GGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCATTTGTTGCTCCTC 1153
 Db |||||
 464 GGTGGCTTAAGACTCAGTGTCTGAGCATGGAATCTCCACAGAAATGCTATGTTGTTCCAC 523
 Qy 1154 ACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTCACGTGCAAG 1213
 Db |||||
 524 ACTACAATCTCAATGCAACACAGCATTAATTTGGCATTTGAATGACGTGCAAGGCTACAG 583
 Qy 1214 TCGTGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTGTGTTG 1273
 Db |||||
 584 TAGTGAATGCAATGCGCAACACCTGTGTTCGATGAGGAATCTAGAAGCTGGCGGCTGATTGA 643
 Qy 1274 TGGTGCCACAGAACTTTCGCGCTGCGTGGAAAGTCCGACAGCGAGAACTTTCGAATACGTGG 1333
 Db |||||
 644 TCGTGCCACAAACTTTTGTGTTGACGCAAAATCAGTGAAGGACAGGTTTCACTTATGTTT 703
 Qy 1334 CATTCAAGACAGACTCAAGGCCCAGCATAGCCAACTCGCCGGTGAACCTCCGCTCATAG 1393
 Db |||||
 704 CATTCAAGACCAATGATTAATGCGCAATTTGCCAGGCTTGCGAGGACACAATTCACCTCTAA 763
 Qy 1394 ATAACTGCGGAGGAGGTGGTTGCAAAATTCATATGCGCTTCCAAAGGAGCAGGCAAGGC 1453
 Db |||||
 764 GTGGTATGCCAGTGGATGTGCTTGCAGCTACATTTCAACATGGACAGGAATGAGGCAAGGC 823
 Qy 1454 AGCTTAAGAACAAACACCCCTTCAAGTTCCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGG 1513
 Db |||||
 824 AGCTCAAAACCAACAATCTCTTTTAATTTCTAGTTCACCCCGTGAGTCCGAAACGACAG 883
 Qy 1514 CTGTGGCTTA 1523
 Db |||||

Db	884	CTCAGCTTA 893	
RESULT 13			
BQ123283			
LOCUS		974 bp mRNA linear	EST 17-APR-2002
DEFINITION		GLSD Medicago truncatula cDNA clone pGLSD-31N1, mRNA	
ACCESSION	BQ123283		
VERSION	BQ123283.1	GI:20175245	
KEYWORDS	EST:		
SOURCE	Medicago truncatula (barrel medic)		
ORGANISM	Medicago truncatula		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
COMMENT	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.		
REFERENCE	1 (bases 1 to 974)		
AUTHORS	Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T.,		
	Cheung, F., and Fraser, C.M.		
	ESTs from late stage developing seeds of Medicago truncatula		
	Unpublished (2002)		
	Contact: Grusak, M.A.		
	USDA/ARS Children's Nutrition Research Center		
	Baylor College of Medicine		
	1100 Bates Street, Houston, TX 77030-2600, USA		
	Tel: 713 798 7044		
	Fax: 713 798 7078		
	Email: mgrusak@bcm.tmc.edu		
	TIGR sequence name: MTRAO73TK More information is available at:		
	www.medicago.org		
	Seq primer: SKmod (CTA GAA CTA gtg gat CC).		
FEATURES	Location/Qualifiers		
source	1..974		
	/organism="Medicago truncatula"		
	/mol_type="mRNA"		
	/cultivar="A17"		
	/db_xref="taxon:3880"		
	/clone="pGLSD-31N1"		
	/tissue_type="Immature seeds"		
	/dev_stage="25 to 35 days after pollination"		
	/lab_host="XLOLR"		
	/clone_lib="GLSD"		
	/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:		
	XhoI; Immature seeds, collected from pods ranging in age		
	from 25 to 35 days after pollination, were harvested from		
	greenhouse-grown plants. Seed were removed and separated		
	from pod walls and immediately frozen in liquid nitrogen.		
	Seeds throughout the age range were pooled for mRNA		
	extraction. cDNA was prepared from polyA+ enriched RNA.		
	The cDNA was directionally ligated into the Unizap XR		
	vector from Stratagene and packaged using Gigapack III		
	Gold packaging extracts. Plasmids containing cDNA inserts		
	were excised from the recombinant lambda-Zap phage using		
	Ex-assist helper phage and propagated in XLOLR cells."		
ORIGIN			
Query Match	21.1%;	Score 321.6;	DB 5; Length 974;
Best Local Similarity	63.7%;	Pred. NO. 3e-77;	
Matches 524;	Conservative 0;	Mismatches 289;	Indels 9; Gaps 2;
QY	680	GAGCAGCAAGAAGAAAGAAAGGTTGGAAACATCTTTCAGCGGCTTCAGCCGGAGT	739
Db	102	GAGGAGGGNACACAGAAATAGAGAGGGCAACATTTTCAGTGGCTTCAGAGGGATT	161
QY	740	TCTTGAAACAAGCTTCCAGTTGACGACAGACAGATAGTGCAAAACCTTAAGAGGCGAGA	799
Db	162	TCTTGAAGATGCAATGAACGT---GAACAGGCATATAGTTGAAAACCTTCAAGGCAGGA	218
QY	800	CCGAGAGTGAAGAGAGGGAGCCATTGTGACAGTGAGGGGAGCCCTCAGATCTTGAGCC	859
Db	219	ATGAAGACCAGGAGAGGGAGCCATTGTCAAGTGAAGAGGACTCAGCATCATGAGCC	278
QY	860	C-----AGATAGAAAGAGACGTCGCCGACGAAGAAGAGGAATACGATGAAGATGAATATG	913
Db	279	CTCCAGAGAGACAACAACGCCACCCAGCAGACAGAAGATGAAGATGAAGAGG	338
QY	914	AATACGATGAAGAGATAGAAGCGTGGCAGGGGAAGCAGAGCGGGGAATGGTATTG	973
Db	339	ATGAGTGGAGGCCACATCATCAGAAAGCAGAAAGAAAGAAAGGCAACAGCGTCTTG	398
QY	974	AAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTCGTAGAAACAGATCCCTTGACA	1033
Db	399	CGGAAACAATTTGACTGCCAGGCTTCCAGAGAACATGCTCTCATCTTCATCACCTGACA	458
QY	1034	TCTACAACCTCAAGCTGTTCACTCAAACTGCCCAACGATCTCAACCTTCTTAATCTTA	1093
Db	459	TCTACAACCTCAAGCTGTAGAAATCAAACTATCACAGCTTCGACCTCCCGAGCTCTCA	518
QY	1094	GGTGGCTTGAGACTAGTGTGAATATGGAATCTCTACAGGAATGCATTTGTTGCTGCTC	1153
Db	519	GGTGGCTAAGACTCAGTGTGAGCATGGAACCTCTCCACAGAAATGCTATGTTGTTCCAC	578
QY	1154	ACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGAGGGCTCAGCTGCAAG	1213
Db	579	ACTACAATCTCAATGCAACAGCAATAATTTGGCATTTGAATGGACGTGCAAGGTACAAG	638
QY	1214	TCGTGACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTCTTG	1273
Db	639	TAGTGAACCTGCAATGCAACACTGTGTTTCGATGAGGAACCTAGAGCTGCCGTCATTGA	698
QY	1274	TGGTGCACAGAACTTTCGCCGTCGCTGGAAAGTCCAGAGCGAGAACTTTCGGAATACGTGG	1333
Db	699	TGCTGCCACAAACCTTGTCTTTCGAGCAAAAATCAGTGAGCGACAGGTTTCACTTATGTTT	758
QY	1334	CATTCAAGACACACTCAAGGCCAGCATAGCAACCTCCCGGTGAAACTCCGTCTATAG	1393
Db	759	CATTCAAGACCAATGATATGCCGAATTCGACGGCTTCAGGGACACANTCCACTCTAA	818
QY	1394	ATAACCTGCCGAGGAGGTGTTGCAAAATTCATATGGCTCCAAAGGGAGCAGGCAAGGC	1453
Db	819	GTGGTATGCCAGTGGATGTCTTCAGCTTACATTCACATGACAGAGGAATGAGGCGAGC	878
QY	1454	AGCTTAAGAACAAACCCCTTCAAGTCTTCTGTTCCACCGT	1495
Db	879	AGCTCAAAAACCAACATCTCTTTAATTTCTAGTTCAACCCCGT	920
RESULT 14			
CA857975			
LOCUS		886 bp mRNA linear	EST 18-DEC-2002
DEFINITION		GLSD Medicago truncatula cDNA clone pGLSD-26J23, mRNA	
ACCESSION	CA857975		
VERSION	CA857975.1	GI:27232339	
KEYWORDS	EST:		
SOURCE	Medicago truncatula (barrel medic)		
ORGANISM	Medicago truncatula		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
COMMENT	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.		
REFERENCE	1 (bases 1 to 886)		
AUTHORS	Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T.,		
	Cheung, F., and Fraser, C.M.		
	ESTs from late stage developing seeds of Medicago truncatula		
	Unpublished (2002)		
	Contact: Grusak, M.A.		
	USDA/ARS Children's Nutrition Research Center		
	Baylor College of Medicine		
	1100 Bates Street, Houston, TX 77030-2600, USA		
	Tel: 713 798 7044		
	Fax: 713 798 7078		
	Email: mgrusak@bcm.tmc.edu		
	TIGR sequence name: MTRW60TK		
	More information is available at: www.medicago.org		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 250.731 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcagcaaccggaggagaa.....ctccgaggctgtggttaa 1524

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517.4	34.0	1446	3	US-09-108-010B-13
2	517.4	34.0	1446	4	US-09-758-652-13
3	517.4	34.0	1446	4	US-10-684-651-13
4	513	33.7	1488	3	US-09-108-010B-11
5	513	33.7	1488	4	US-09-758-652-11
6	513	33.7	1488	4	US-10-684-651-11
7	513	33.7	1743	4	US-09-762-381-1
8	504	33.1	1458	3	US-09-108-010B-12
9	504	33.1	1458	4	US-09-758-652-12
10	504	33.1	1458	4	US-10-684-651-12
11	198	13.0	1551	3	US-09-108-010B-15
12	198	13.0	1551	4	US-09-758-652-15
13	198	13.0	1551	4	US-10-684-651-15
14	153.2	10.1	1689	3	US-09-108-010B-14
15	153.2	10.1	1689	4	US-09-758-652-14
16	153.2	10.1	1689	4	US-10-684-651-14
17	142	9.3	1706	4	US-09-462-720-1
18	136.6	9.0	1685	1	US-08-486-721A-1
19	132.2	8.7	1556	1	US-08-486-721A-2
20	114.6	7.5	3113	1	US-08-146-422-20
21	114.6	7.5	3113	1	US-08-626-554-2
22	108.4	7.1	4939	4	US-09-645-593-8
23	93.8	6.2	7218	1	US-08-232-463-14
24	53.4	3.5	3477	4	US-09-462-720-3
25	49.8	3.3	15252	4	US-09-949-016-13584
26	49.4	3.2	150394	4	US-09-949-016-13042
27	47.2	3.1	289	3	US-09-007-005-17

28	47.2	3.1	289	3	US-09-244-796-17	Sequence 17, Appl
29	46.8	3.1	247781	4	US-09-949-016-14193	Sequence 14193, A
30	46.8	3.1	276687	4	US-09-949-016-13840	Sequence 13840, A
31	46.4	3.0	5394	3	US-08-688-376-1	Sequence 1, Appli
32	46.4	3.0	10409	3	US-08-772-440-33	Sequence 33, Appl
33	44.8	2.9	165651	4	US-09-949-016-13032	Sequence 13032, A
34	44.4	2.9	929	4	US-09-671-317-14	Sequence 14, Appl
35	44.4	2.9	1001	4	US-09-671-317-439	Sequence 439, App
36	44.4	2.9	2277	1	US-08-676-967-2	Sequence 2, Appli
37	44.4	2.9	2277	1	US-08-676-974-2	Sequence 2, Appli
38	44.4	2.9	2277	2	US-09-949-016-14543	Sequence 14543, A
39	43.2	2.8	387902	4	US-09-949-016-12557	Sequence 12557, A
40	43.2	2.8	421883	3	US-08-658-136-2	Sequence 2, Appli
41	42.4	2.8	53526	3	US-08-658-136-2	Sequence 1, Appli
42	42.4	2.8	53577	3	US-08-658-136-2	Sequence 1, Appli
43	42.2	2.8	29927	4	US-09-949-016-17474	Sequence 17474, A
44	42.2	2.8	29927	4	US-09-949-016-17474	Sequence 17474, A
45	42.2	2.8	29927	4	US-09-949-016-17475	Sequence 17475, A

ALIGNMENTS

RESULT 1
US-09-108-010B-13
; Sequence 13, Application US/09108010B
; Patent No. 6362399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION NUMBER: US/09/108,010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNN M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-108-010B-13

Query Match 34.0%; Score 517.4; DB 3; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.1e-144;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

```
QY 5 AGCAACCGGAGGAGAACCGCTGCGAGTTCCAGGCGCTCAATGCGCGAGAGACCTGACAATC 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 65 GCATTGAATCAGAGGGGGTTACATTGAGACTTTGGAAACCCCAACACACAGGAGTTGCAAT 124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 GTATAGAGTCAGAAAGTGGCTTCATTGAGACATGGAACCCCTAACAAAGCCATTCCAGT 190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 GCCCGGCGTCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCTTCT 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 GTCCCGGTGTGCCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTCGCGAGACCTTCCT 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 185 ACTTCCAAATGCTCCCGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTCATAT 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 ACACCAACGCTCCCGAGGAGATCTACATCCAAAGGTAGTGATTTTGGCATGATAT 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 TCCTCGTTGTCTAGACACTATGAAGAGGCTCACACAAAGTCTCGATCTCAGTCCC 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 TCCCGGTTGTCTTAGCACATTTGAAGAGCTT----- 342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 AAAGACCACCAAGACGTCTCCAGGAGAGAACCAAGCCAAACAGCAACGAGATAGTCACC 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 343 -----CAACAAAGGACAAAGCAGACGAGCCCAAGACCGTCACC 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 365 AGAAGGTGCAACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCACCGGTGTTGCTTCT 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTGCAGTGCCAAACCGGTTTGCATACT 442
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 425 GGCTCTAACACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACAA 484
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 443 GGATGTACAAACAAATGAAGACACTCTCTGTTGTTGCCGTTTCTCTTATTGACACCAACAGCT 502
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 485 ACCACAACAGCTTGATCAGTCTCCAGGAGATTCATTTGCGTGGGAACACGAGCAAG 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 503 TCAGAACCAAGCTCGACAGATGCTTAGGAGATCTATCTTGCTGGGAACCAAGAGCAAG 562
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 545 AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 563 AGTTCTACATGATCAGCCACAGACGACGAAG----- 595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 CATACAGCCGCAAGTCAAGCTAGACAAAGAGCGTGAAATTAGCCCTCGAGGACAGC 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 596 -----GAGGTACTC 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 665 ACAGCCGAGAGAACGAGCAGGACAAAGAAAGAAACGAAGTGAACATCTTCAGCG 724
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 AAAGCCAGAAAGGAAGCGTCACCAAGAGAGAAACGAAGGAGCAGCATATTGAGTG 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 725 GCTTCAACGCGGAGTTCTTGGAACAAAGCCTTCCAGGTTGACGACACAGATAGTCNAA 784
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 665 GCTTCGCCCGGAATCTTGGAACATGCGTTT---GTGCTGGACAGCGAGATGAGAA 721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 785 ACCTAAGAGCGAGACGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGGAGGCC 844
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 722 AGCTACAAGGTGAGAACGAAGAGAGAGAGAGAGAGGTTGCTCATTTGTGACAGTGAAGAGGATC 781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 845 TCAGAACTTTGAGCCAGATAGAAAGAGAGAGCTGCCGACGAAGAGAGGAATACATGAAG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 782 TCAGCGTGATAGCCCAACCCACGGAAGAGCAGCAACAAAGACCCGAGGAGAGGAAGC 841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 905 ATGAATATGAATACGATGAAGAGGATAGAAGGCGTGGCGAGGGAAGCAGAGCGAGGGA 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 842 CAGATTGTGACGAGAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 965 ATGGTATTGAAGAGAGATCTGACCGCAAGTGCTAAAAAGAACATTTGGTAGAAACAGAT 1024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 887 ATGGCATTTGACGAGACCAATTGACAAATGAGACTTCGCCACACACATTTGCCAGACTTCAT 946
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1025 CCCTGACATCTCAACCCCTCAAGCTGGTTCTCACTCAAAAACGCAACGATCTCAACCTTC 1084
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCCGTACAGCTCGACTTC 1006
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 1085 TAATACTTAGTGGCTTGGACCTAGTGTGAAATATGGAATCTCTACAGGAATGCAATTGT 1144
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1007 CAGCCCTCTCGTGGCTCAAACCTCAGTGGCCAGTTTGGATCACTCCGCAAGAATGCTATGT 1066
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1145 TTGTGCTCACTACAACACCAACGACACAGCATCATATATCGATTGAGGGACGGGCTC 1204
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1067 TCGTGCCACACTACAACCTGAAACGCAACAGACATAATATACGCATTTGAATGACCGGCAT 1126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1205 AGGTGCAAGTCTGTGACAGCAGCAACGAGAGTGTACGACGAGGAGCTTCAAGAGGGTC 1264
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1127 TGTACAAGTGTGATTTGCAATGTTGAGAGAGTGTGATGAGAGAGCTGCAAGAGGGAC 1186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1265 AGTGTCTTGTGTGCCACAGAACTTTCGCGTTCGCTGGAAGTCCAGAGCGAGAACTTCG 1324
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1187 AGGTGTTAATTGTGCCACAAACCTTTCGCGTGGCTGCAAGATCACAGAGCGCAACTTCG 1246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1325 AATAGTGGCATTTCAAGACAGACTCAAGCCCGACAGTACCAACCTTCGCCGCTGAAACT 1384
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1247 AGTATGTTTCAATTCAAGACCAATGATAGACCTTCGATCGGCAACCTTCAGGTGCAACT 1306
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1385 CCGTCATAGATAACCTGCGGAGGAGGTGTTGCAAAATTCATATGGCCTCCAAAGGGAGC 1444
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1307 CATTGTTGACGCATTTGCGGAGGAGTGATTCAGCAAACTTTTAACCTTAAGAGGCGC 1366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1445 AGCAAGCAGCTTAAGAACAAACACCCCTTCAAGTTCCTTCGTTCCACCGTCTCAGAGT 1504
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1367 AGGCCAGGCGAGTCAAGAACAAACACCTTTCAGCTTCTCGTTCCACCTCAAGGAGTCTC 1426
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1505 CTCGAGGGCTGTGGCTTA 1523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1427 AGAGGAGAGTTGTGGCTTA 1445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-09-758-652-13
; Sequence 13, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; ;
```

[illegible]

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/108,010B
FILING DATE: 30-Jun-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14, 1996
ATTORNEY/AGENT INFORMATION:
NAME: LYNNE M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1488 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-108-010B-11

Query Match 33.7%; Score 513; DB 3; Length 1488;
Best Local Similarity 62.7%; Pred. No. 2.3e-143;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;

Qy 5 AGCAACCGGAGGAAACGGTGGCGAGTTCAGGCGCTCAATGCGCAGAGACCTGACAAATC 64
Db 71 AGCAGCTTCAGAAAACGAGTGGCGAGATCCAAAACCTCAATGCCCTCAAAACCGGATAAC 130

Qy 65 GCATTGATCAGAGGGGGGTACATTGAGACTTGGAAACCCCAACACAGAGGTTGCAAT 124
Db 131 GTATAGAGTCAGAGGAGGGCTCATTTGAGACATGGAAACCCCAACAAAGCCATTCCAGT 190

Qy 125 GCGCGCGGTGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCGTTTCT 184
Db 191 GTGCGGTGTGCGCTCTCTCGCTGACCTCAACCGCAACGCCCTTCGTAGACCTTCT 250

Qy 185 ACTCCAAATGTCGCCAGGAGATCTTCATCCAGCAAGGAGGGATATCTTTGGGTGATAT 244
Db 251 ACACCAACGTTCCCGAGGAAATCTACATCCAAAGGTAAGGGTATTTTTGGCATGATAT 310

Qy 245 TCCTGTGTTCTAGACACTATGAAGAGCTTCACACAAAGGTGCTGATCTCAGTCCC 304
Db 311 ACCCGGTTGTCTAGACACTATTTGAAGAGCTCA----- 344

Qy 305 AAAGACCAACAGAGCTCTCAAGGAGAGCAACAAAGCCCAACAGCAGAGATAGTCAAC 364
Db 345 -----ACAACCTCAACAAGAGAGCAAGAGCAGACACCAAGCCGTCACC 391

Qy 365 AGAAGGTGCAACCGTTTCGATGAGGGTATCTCATTTGAGTTCGCCACGGTGTGCTTTCT 424
Db 392 AGAAGATCTATACTTCAGAGAGGGTGTATTTGATCGCAGTGCCTACTGTTGGTGTGAT 451

Qy 425 GGCTCTACAGGACCAAGCACTGATGTTGTTGCTGTTTCTCTTACTGACCAACACACA 484
Db 452 GGATGTACAAATGAAGACACTCTGTTGTTGCGGTTTCTATTTATTTGACACCAACAGCT 511

Qy 485 ACAGCAACCAAGCTTGATGAGTTCGCCAGGAGATTCAAATTTGGTGGGAAACACCGGCAAG 544
Db 512 TGAGAACCAAGCTTCGACCAAGATGCTTAGGAGATCTTATCTTGTGGGAACCAAGAGCAAG 571

Qy 545 AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGAGCAAGCAAGCAAGCTTACCATATAGCC 604
Db 572 AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGAGCAAGCAAGCAAGCAAGCAAGCAAG 594

Qy 605 CATAAGCCCGCAAGTTCAGCTAGACAAAGAGCGTGAATTTAGCCCTTCGAGGACGC 664
Db 595 -----CAAGGAGGTCAATC 607

Qy 665 ACAGCCGCGAGAGAAACGAGCAGGAGCAAGAAAGAAACGAAAGTGGAAACATCTTCAGCG 724
Db 608 AAAGCCAGAAAGGAAAGCATCAGCAGAAAGAAAGAAAGGAGGAGCAGATATTGAGTG 667

Qy 725 GCTTCAAGCCGAGGATTCCTGGAACAAAGCCTTCCAGGTTGACGACAGATAGTGCAAA 784
Db 668 GCTTCAAGCCCTGGAATCTTGGAAACATGCTTTCAGCT---GGCAAGCAGATAGCGGAAA 724

Qy 785 ACCTAAGAGCGCAGACGAGAGTGAAGAGAGGAGGACCATTTGACAGTTCAGGAGGAGCC 844
Db 725 ACCTAAGAGAGAGAAACGAGGAGGAGCAAGAGGAGGACCATTTGACAGTGAAGAGGAGTC 784

Qy 845 TCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCGTCCGACGAGAAAGAGAGAAAT 895
Db 785 TGAGCGTGATTAACCAACCCACCGAGCAGCAGCAACAAAGACCCAGGAGAGAGAGAGAG 844

Qy 896 ACATGAAGATGAATA-----TGAATCAGATGAAGAGGATAGAGAGCGGTG 940
Db 845 AAGAGAGAGTGAAGAGCCACAGTGCAGAGGTAAGAGCAAACTGCCAACGCCCGCAG 904

Qy 941 GCAGGGAAGCAGAGGAGGAGGGAATGTTATTAAGAGAGCATCTCGACCGCAAGTGCTA 1000
Db 905 GAAGCCAAAGCAAGAGCAAGAAATGGCATTTGACGAGCATTATGACCATGAGACTTC 964

Qy 1001 AAAGAAACATTTGTTAGAAAACAGATCCCTGACATCTTACAAACCTCAAGTGGTTCACTCA 1060
Db 965 GCCACAACATTTGGCCAGACTTCATCACTGACATCTTACAACTTCAAGCGGTAGGCTCA 1024

Qy 1061 AAATCCCAACGATCTCAACCTTCTAATACTTAGTGGTGGTGGACCTAGTGTGAATATG 1120
Db 1025 CAACCCCAACGAGCTTGAATCTTCCAGCCCTCTCGTGGTTCAGACTCAGTGTGAGTTG 1084

Qy 1121 GAAATCTCTACAGGAGTGAATGTTGTTGCTGCTCACTACAAACCCAGCAGCAGCATCA 1180
Db 1085 GATCTCTCCGCAAGATGCAATGTTGCTGCCACACTTACAACTGAAACGCGAACAGCATAA 1144

Qy 1181 TATATCGATTGAGGGAGCGGGTCACTGCAAGTCTGTGACAGCAACGCGCAACAGAGTGT 1240
Db 1145 TATACCATTTGATGACCGGGCATTCATACAGTGTGATTTGCAACGGTGAGAGAGTGT 1204

Qy 1241 AGCAGAGAGCTTCAAGAGGGTCACTGCTTGTGTTGTCACAGAACTTTCGCGTGCCTG 1300
Db 1205 TTGATGGAGAGCTGCAAGAGGGACGGTGTGATCTGTCGCAAAAACCTTGTGGTGGCTG 1264

Qy 1301 GAAAGTCCAGAGCGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCCGCA 1360
Db 1265 CAAGATCAAGAGTGAACCTTCGAGTATGTTGATTCAGGACCAATGATACACCATGA 1324

Qy 1361 TAGCCAACTCGCGGTGAAACTCCGTTCATAGATAAAGTCCGCGAGGAGGTGGTTGCAA 1420
Db 1325 TCGGCACTCTTGACGGGGCAAACTCATTTGTAACCATTTACAGAGGAGAGTATTCAGC 1384

Qy 1421 ATTATATATGCTTCAAGAGGAGCAGGCAAGGAGCTTAAAGCAACAAACCCCTTCAAGT 1480
Db 1385 ACATTTCAACCTTAAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1444

Qy 1481 TCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTA 1523
Db 1445 TCTGTTTCCACTCAGGAGTCTCAGAGAGAGAGTGTGGCTTA 1487

RESULT 5
US-09-758-652-11
; Sequence 11, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

/ CLASSES OF SOYBEAN SEED
/ PROTEIN GENES
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
/ STREET: 1007 MARKET STREET
/ CITY: WILMINGTON
/ STATE: DELAWARE
/ COUNTRY: UNITED STATES OF AMERICA
/ ZIP: 19898
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: DISKETTE, 3.50 INCH
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: MICROSOFT WINDOWS 95
/ SOFTWARE: MICROSOFT WORD 7.0A
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/758,652
/ FILING DATE: 11-Jan-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/019,940
/ FILING DATE: JUNE 14,1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: LYNNE M. CHRISTENBURY
/ REGISTRATION NUMBER: 30,971
/ REFERENCE/DOCKET NUMBER: BB-1071-A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 302-992-5481
/ TELEFAX: 302-773-0164
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1488 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-758-652-11

Query Match 33.7%; Score 513; DB 4; Length 1488;
Best Local Similarity 62.7%; Pred. No. 2.3e-143;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;
QY 5 AGCAACGGAGGAGAACGGGTGCGAGTCCAGCGCTCAATGGCGAGAGACCTGACAATC 64
DB 71 AGCAGCTCAGCAAAACGAGTGCAGATCCAAAACCTCAATGCGCTCAAAACCGGATAAC 130
QY 65 GCATTGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACAAACCGAGGTTCGAAT 124
DB 131 GTATAGAGTCAGAGGAGGCGCTCATTTGAGACATGGAACCCCAACAAAGCCATTCCAGT 190
QY 125 GCGCCGGCTGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCTTTCT 184
DB 191 GTCCCGGTGTTCCTCTCTCGCTGACCTCAACCGCAACGCCCTTCGTAGACCTTCCT 250
QY 185 ACTCCAAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAAAGGGATATCTTTGGGTTCATAT 244
DB 251 ACACCAACGCTCCAGGAAATCTACATCAACAAAGTAAGGTATTTTGGCATGATAT 310
QY 245 TCCTCGTGTCTTAGACACTATGAAAGAGCTTCAACAGAGCTTCAGTCCATCTCAGTCCC 304
DB 311 ACCCGGGTGTCTTAGACATTTTGAAGAGCTCA----- 344
QY 305 AAAGACCACCAAGACGCTCTCCAGGAGNAGACCAAGCCCAACAGCAACGAGATGTCACC 364
DB 345 -----ACAACCTCAACAAAGAGGACAAAAGCAGCAGACCAACGACCGTCACC 391
QY 365 AGAAGGTGCACCGCTTTTCGATGAGGGGTGATCTCATTCAGTTCACCAACCGGTGTTGCTTCT 424
DB 392 AGAAGATCTATACTTCAGAGAGGGGTGATTTGATCGAGTGCCTACTCTGTTGTCATGCT 451
QY 425 GGCTCTACAAACGACCAAGACACTGATGTTGTTGCTGTTCTTCTTCTGACACCAACAA 484

DB 452 GGATGTACAAACAATGAAGACACTCTCTGTTGTTGCGGTTTCTATTATTGACACCAACAGCT 511
QY 485 ACGACAACCCAGCTTGATCAGTTCCCGAGAGATTCAATTTTGGCTGGACACCGAGCAAG 544
DB 512 TGGAGAACCCAGCTCGACCCAGATGCGCTTAGGAGATTCTATCTTCTGGGAACCAAGAGCAAG 571
QY 545 AGTTCTTAAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTTACCATATAGCC 604
DB 572 AGTTCTTAAATAATCAGCAAGAG----- 594
QY 605 CATACAGCCCGCAAAAGTTCAGCCTTAGACAAGAAGAGCGTGGAATTTAGCCCTTCAGAGCAGC 664
DB 595 -----CAAGGAGGTTCATC 607
QY 665 ACAGCCGACAGACGACGAGCAGCAGCAAGAGAAAGAAAACGAAGGTGGAACATCTTTCAGCG 724
DB 608 AAAGCCAGAAAGGAAGCATCAGCAAGAAAGAAAGAAAACGAGGAGCGACATATTGAGTG 667
QY 725 GCTTCACGCGGAGTTCTCTGGAACCAAGCCTTCCAGGTTTGACGACAGACAGATAGTGCAAA 784
DB 668 GCTTCACCCCTGGAATTTCTTGGAAACATGCATTCAGCGT---GGACAAGCAGATAGCGAAA 724
QY 785 ACCTAAGAGCGAGACCGAGAGTGAAGAGAGGGAGGCCATTTGTGACAGTGAAGGAGGCC 844
DB 725 ACCTAAGAGGAGAAACGAAGGGGAAGACAAGGGAGCCATTGTGACAGTGAAGAGGAGTTC 784
QY 845 TCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCGACGAAAGAGAGGAAT 895
DB 785 TGAGCGTGATAAAACCCACGAGCAGACGACCAACAAGACCCTCCAGNAGAGGAGAAAG 844
QY 896 ACGATGAAGATGAATA-----TGAATACGATGAAGAGGATGAAGAGCGGTG 940
DB 845 AAGAAGAGGATGAGAAGCCACAGTCAAGGGTAAAGACAACACACTGCCAACCCCGGAG 904
QY 941 GCAGGGGAAGCAGAGCGCAGGGGGAATGTTATGAAGAGAGCATCTGCACCGCAAGTGCTA 1000
DB 905 GAAGCCAAAGCAAAAGCAGAAATGTCATTTGACGAGACCAATATGACCATGAGACTTC 964
QY 1001 AAAAGAACAATTCGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGTCTCACTCA 1060
DB 965 GCCACAACATTTGGCAGACTTCATCACCCTGACATCTCAAAACCTCAAGCCGTTAGGCTCA 1024
QY 1061 AAACCTGCCAACGATCTCAACCTTCTAATACTTAGGTGGCTTGAGCTAGTGTGTAATATG 1120
DB 1025 CAACCGCCACCAAGCTTGACTTCCAGGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTG 1084
QY 1121 GAAATCTCTAAGGAATGCAATTTGTTGCTCCTACTACAAACCAACGACGACAGCATCA 1180
DB 1085 GATCTCTCGCAAGAATGCAATGTTCTGTCGCCACACTACAAACCTGAAACGCAACAGCATAA 1144
QY 1181 TATATCGATTGAGGGACGCGGCTCAGTGCMAAGTCTGAGACAGCAACGGCAACAGAGTGT 1240
DB 1145 TATACGCAATGATGAGCGGGCATTTGATACAGTGTGTAATTTGCAACGGTGAGAGATGT 1204
QY 1241 ACGACGAGAGCTTCAAGAGGTCAGTCTTGTGTCGACAGAACTTTCGCGCTGCTG 1300
DB 1205 TTGATGGAGAGCTGCAAGAGGGACGGGTCTGATGCTGTCACAAAACCTTGTGTGGCTG 1264
QY 1301 GAAAGTCCAGAGCAGAGACTTCGAAATACGTGGCAATCAAGACAGACTCAAGGCCCGACA 1360
DB 1265 CAAGATCAGAGTGACAACTTCGAGTATGTGTCTTCAAGACCAATGATCACCCATGA 1324
QY 1361 TAGCCAAACCTCCCGGTGAAAACTCCGTCATAGATAAAGCTGCCGGAGGAGGTGTTGCAA 1420
DB 1325 TCGGCACTCTTCAGGGGCAAACTCATTTGTGAAACGCATTACAGAGGAGATGATTCAGC 1384
QY 1421 ATTTCATATGGCTCCAAAGGGAGCAGGCAAGCAGCTTAAGAAACAAACACCTTCAAGT 1480
DB 1385 ACACCTTCAACCTTAAAGAGCCAGAGGCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444
QY 1481 TCTTGTCTCCACCGCTCTCAGCAGTCTCCGAGGGCTGTGCTTTA 1523
DB 1445 TCCTGTTTCCACCTCAGGAGTCTCAGAAAGAGAGCTGTGCTTTA 1487

RESULT 6
US-10-684-651-11
; Sequence 11, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-684-651-11

Query Match 33.7%; Score 513; DB 4; Length 1488;
Best Local Similarity 62.7%; Pred. No. 2.3e-143;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;
QY 5 AGCAACGGAGGAGAACCGCTGCGAGTTCCAGCGCTCAATGCGCAGAGACCTGACATC 64
DB 71 AGCAGCCTCAGCAAAACGAGTGGCAGATCAAAAACCTCAATGGCCCTCAAAACCGGATAACC 130
QY 65 GCATTGAATCAGAGGGCGGTTACATTGAGACTTGGACCCCAACACACAGGAGTTCGAT 124
DB 131 GTATAGAGTCAGAGGAGGGCTCATTTGAGACATGGAACCCCTAAACAAAGCCATTCCAGT 190
QY 125 GCGCGCGCTGCGCCCTCTCTCGTTAGTCTCGCGCCCAACGCGCTTCGTAGGCGTTTCT 184
DB 191 GTGCGGCTGTGCGCCCTCTCTCGTGCACCTCAACCGCAAGCCCTTCGTAGACCTTCT 250
QY 185 ACTCCAAATGCTCCCGCAGGAGATCTTCATCCAGCAAGGAAGGAGATCTTTGGGTGATAT 244
DB 251 ACACCAACGGTCCCGCAGGAATCTACATCCAAAGTAAGGGTATTTTGGCATGATAT 310
QY 245 TCCTGGTCTCTAGACATATGAGAGCGCTCACACAAAGGTCGTGCTCAGTCTCAGTCCC 304

DB 311 ACCGGGTTGTCTCTAGCACATTTGAAGAGCCTCA----- 344
QY 305 AAAGACCAACAAGACGCTCTCCAAAGGAGAGAACAAAGCCAAACAGCAACGAGATAGTCACC 364
DB 345 -----ACAACTCTCAAAAGAGGACAAAGAGCAGCAGACCAAGACCGTCACC 391
QY 365 AGRAGGTGCACCGTTTTCGATGAGGTGATCTCATTTGCTGCTGCTTCTTACTGACCAACCAACA 424
DB 392 AGAAGATCTATAACTTCAGAGAGGGTGATTTGATCGAGTGCTCTACTGGTGTGCTATGGT 451
QY 425 GGCTCTACAAACGACCAACGACACTGATGTTGTTGCTGCTGCTTCTTACTGACCAACCAACA 484
DB 452 GGATGTACAACAATGAAGACACTCTCTGTTGTTGCTGCTTCTTACTGACCAACCAACAGCT 511
QY 485 ACAGCAACAGCTTGTATGATGATGTTCCCGAGAGATTCATTTGGCTGGGAAACAGGAGCAAG 544
DB 512 TGGAGAACAGCTCGACCAAGATGCTCTAGGAGATTTCTATCTGCTGGGAAACAGAGCAAG 571
QY 545 AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCATATAGCC 604
DB 572 AGTTCTTAAATATATCAGCAAGAG----- 594
QY 605 CATACAGCCCGCAAAAGTCAGCCTTAGACAAGAGAGCGTGAAATTTAGCCCTTCGAGGACAGC 664
DB 595 -----CAAGAGGTATC 607
QY 665 ACAGCCGACAGAGAACGAGCAGGACAAAGAAAGAAACGAGGTGAAACATCTTCAGCG 724
DB 608 AAAGCCAGAAAGGAAAGCATCAGCAAGAAAGAAACGAGGAGCAGCATATTGAGTG 667
QY 725 GCTTCAGCGCGGAGTTCTTGGAAACAGCTTCCAGGTTGACGACACACAGATAGTCGAAA 784
DB 668 GCTTCACTCCCTGGAAATCTTGGAAACATGCACTTCAGCGT---GGCAACAGCAGATAGCGAAA 724
QY 785 ACCTAAGAGCGAGACCGGAGGTGAAGAGAGAGGAGCCATTTGTGACAGTGTGAGGCGAGGCC 844
DB 725 ACCTAAGAGAGAGACGAGGAGGAGAGCAAGGAGCCATTTGTGACAGTGAAGAGAGGTC 784
QY 845 TCAGAACTTTGAGCCCA-----GATGAAAGAGAGCGTCCGACGAGAGAGAGAAAT 895
DB 785 TGAGCGTGATAAAACCAACCGCAGCAGCAGCAACAAAGACCCAGGAGAGAGAGAAAG 844
QY 896 ACGATGAAGTGNATAT-----TGATACGATGAAGAGGATAGAGCGGTG 940
DB 845 AAGAGAGGATGAGAAAGCCACAGTGCAAGGGTAAAGACAAACACTCCCAACGCCCCCGAG 904
QY 941 GCAGGGAAAGCAGAGCGGAGGAGTGTATTGAAGAGACGATCTGCACCGCAAGTGCTA 1000
DB 905 GAAGCCAAAGCAAAAGCAGAGAAATGGCATTTGACGAGACCATATGCCACATGAGACTTC 964
QY 1001 AAAAGAAACATTTGGTAGAAAACAGATCCCTGACATCTTAAACCCCTCAAGCTGTTGTTCACTCA 1060
DB 965 GCCAACAACTTTGGCCAGACTTCATCACCTGACATCTTAAACCCCTCAAGCGGTAGCGTCA 1024
QY 1061 AACTGCCAAACGATCTCAACCTTCTAATCTTAGTGGTGGCTTGGACCTAGTCTGATATG 1120
DB 1025 CAACGCCCAACAGCTTGAATCTTCCAGCCCTCTCGTGGTTCAGACTCAGTGTGAGTTG 1084
QY 1121 GAAATCTCTACAGGAATGATTTGTTGCTGCTCACTTAAACCAACCAACGACACAGCATCA 1180
DB 1085 GATCTCTCCGCAAGATGCAATTTGCTGCCCACTTAAACCTGACGACGACGACATAA 1144
QY 1181 TATATCGATTGAGGGAGCGGCTCAGTGTCAAGTCTGTGGACAGCAACGCAACAGAGTGT 1240
DB 1145 TATACGCAATTGAATGGACGGGCAATTGATACAAAGTGTGAAATTCGAACGGTGAGAGAGTGT 1204
QY 1241 ACGACGAGAGCTTCAAGAGGGTCAGTCTTGTGTGTCACAGAACTTCGCGCTGCTG 1300
DB 1205 TTGATGAGAGAGTGCAGAGGAGCGGGTCTGATGCTGTCACAAAACCTTTGTGTGGTGG 1264
QY 1301 GAAAGTCCAGAGCAGAGCACTTCGAAATAGTGTGCAATTCGAAGCAGACTCAAGGCCCAAGCA 1360
DB 1265 CAAGATCAGAGAGTGACAACTTCGAGTATGTGTCTTCAAGCAACCATGATACACCCATGA 1324

Qy	1361	TAGCCAACTCGCGGGTAAAACTCCGTCATAGATAAACCTGCCGAGGAGGTGGTTGCAA	1420
Db	1325	TCGGCACTCTTGTCAGGGGCAAACTACTTGTGTAACGCATTACCAGAGGAAGTGATTCAGC	1384
Qy	1421	ATTCAATATGGCTCTCAAAGGAGCAGCAGCAGCTTAAGAACAAACAAACCCCTTTCAAGT	1480
Db	1385	ACACTTTCAACCTAAAAAGCCAGCAGGCCAGGCAGATATAAGAAACAAACCCCTTTCAAGT	1444
Qy	1481	TCATTGGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTA	1523
Db	1445	TCCTGGTTCCACCTCAGGAGTCTCAGAAGAGAGCTGTGGCTTA	1487

RESULT 7
 US-09-762-381-1
 ; Sequence 1, Application US/09762381
 ; Patent No. 6576820
 ; GENERAL INFORMATION:
 ; APPLICANT: Takaiwa, Fumio
 ; APPLICANT: Utsumi, Shigeru
 ; APPLICANT: Katsube-Tanaka, Tomoyuki
 ; TITLE OF INVENTION: TRANSGENIC PLANT EXPRESSING SOYBEAN GLYCININ
 ; FILE REFERENCE: 201487/1050
 ; CURRENT APPLICATION NUMBER: US/09762,381
 ; CURRENT FILING DATE: 1999-03-04
 ; PRIOR APPLICATION NUMBER: PCT/JP99/01057
 ; PRIOR FILING DATE: 1999-03-04
 ; PRIOR APPLICATION NUMBER: JP 10-223-897
 ; PRIOR FILING DATE: 1998-08-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1743
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (52) .. (1536)
 US-09-762-381-1

Query Match	33.7%	Score	513;	DB	4;	Length	1743;
Best Local Similarity	62.7%;	Pred.	No. 2.5e-143;				
Matches	968;	Conservative	0;	Mismatches	425;	Indels	150; Gaps 5;

Qy	5	AGCAACGGAGGAGAAACCGGTGCCAGTTCCAGCGCCTCAATGCGCAGACCTTGACAATC	64
Db	122	AGCAGCCTCAGCAAACAGTGCAGATCCAAAACCTCAATGCCTCAAACCGGATAACC	181
Qy	65	GCAATTGAATCAGAGGGCGGTTCATTTCAGACTTGGAAACCCCAACCAACGAGAGTTCCGAAT	124
Db	182	GTATAGAGTCNAGAGGAGGGCTCATTCAGACATGGAACCTTAACNACAGCCATTCAGT	241
Qy	125	GGCGCGGTGGCCCTCTCTGGCTTAGTCTTCGCGCGCAACGCGCTTCGTAGGCGTTTCT	184
Db	242	GTGCGGTGTGGCCCTCTCTCGCTGCACCTCAACCGCAACGCGCTTCGTAGACCTTCCT	301
Qy	185	ACTCCAATGTCCTCCGAGAGATCTTCATCCAGCAAGGAAGGGATATCTTTGGGTGTGATAT	244
Db	302	ACACCAACGGTCCCAGGAAATCTACATCCAACCAAGGTAAGGGTATTTTTGGCATGTATAT	361
Qy	245	TCCCTGTGTGCTCTAGACACTATCAAGAGCCTCACACACAAGTCGTCAGTCCC	304
Db	362	ACCGGGTGTGCTTAGACATTTGAGAGCCTCA-----	395
Qy	305	AAAGACCACCAAGACGTTCCAAGGAGAAGCAAAAGCCAACGACAGATAGTCACC	364
Db	396	-----ACAACTCAACAAAGAGGACAAAGCAGCAGACCAACAAGCCGTACC	442
Qy	365	AGAGGTGCACCGTTTCGATGAGGGTGTATCTCATTCGAGTCCCAACGGGTGCTTTCT	424
Db	443	AGAAAGATCTTAACCTTCAGAGAGGGTGAATTGATCGCAGTGCCTACTCGGTGTGTCATGGT	502

Qy	425	GGCTCTAACAAGCAGCAGCAGCACTGATGTTGTGTTGCTGTTCTCTTACTTGACACCAACAACA	484
Db	503	GGATGTACAAACAATAGAGACACTCTCTGTTGTGTCGGCTTTCTATTATTGACACCAACAGCT	562
Qy	485	ACGACACCAACAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTTGGGAACAACGAGACAG	544
Db	563	TGGAGAACCAAGCTCGACAGATGCCATAGAGATTCTATTCTGTCTGGAAACCAAGAGCAG	622
Qy	545	AGTTCTTAAGGTTACCAAGCAACAAGCAGACAACAAGCAGACGAAGAAGCTTACCATATAGCC	604
Db	623	AGTTCTTAATATCAGCAAG-----	645
Qy	605	CATACAGCCCGCAAAAGTCAGCTTAGACAAGAAGAGCGTGAAATTTAGCCCTCGAGGACAGC	664
Db	646	-----CAAGGAGGTCAATC	658
Qy	665	ACAGCCGACAGAACGAGCAGGACAAGAAAGAAACGAAGCTGGAACATCTTCAGCG	724
Db	659	AAAGCCAGAAGAAAGCATCAGCAAGAAGAAAGAAACGAAGGAGGAGCATATTGAGTG	718
Qy	725	GCTTTCAGCCGCGAGTTCCTGGAAACAAGCCCTTCAGAGTTGACGAGACAGACATAGTGCAAA	784
Db	719	GCTTTCACCTTGGAAATCTTGGACATGCAATTCAGCGT---GGACAGCAGATAGCGAAA	775
Qy	785	ACCTAAGAGCGGAGACCGAGAGTGAAGAAGAGGGAGCCATTGTGACAGTGAGGGGAGGCC	844
Db	776	ACCTACAAGGAGAGAACGAAGGGGAAGACAAGGGAGCCATTGTGACAGTGAAGAGGAGGTC	835
Qy	845	TCAGAACTCTTGAGCCCA-----GATAGAAGAGACGTGCGGACGAGAGAGAGGAAT	895
Db	836	TGAGCGGTGATAAACAACCCACCGAGCAGCAGCAACAAGACCCCAAGGAGAGAGAGAAG	895
Qy	896	ACGATGAAGATGAATA-----TGAATACGATGAAGAGGATAGAAGGCGTG	940
Db	896	AAGAAGGATGAGAGCCACAGTGCNAGGTAAAGACAACAACACTGCCAACGCCCGAG	955
Qy	941	GCAGGGGAAGCAGAGCAGGGGAATGGTATTGAAGAGACGATCTGCACGCAAGTGTCTA	1000
Db	956	GAAGCCAAAGCAAAAGCAGAAGAATGGCAATGGCAATGACGAGACCAATGACCATTGAGACTTC	1015
Qy	1001	AAAGAACAATTGTTAGAACAGATCCCTGACATCTACACCTCAAGCTGGTTCCTCA	1060
Db	1016	GCCACAACAATTGGCCAGACTTCATCACTTGACATCTACAACTTCAAGCCGAGGAGCTCA	1075
Qy	1061	AAACTGCCAACGATCTCAACTCTTAATACTTAGTGGCTTGACCTTAGTCTGGAATATG	1120
Db	1076	CAACCGCCACCGCTTGACTTCCAGCCCTCTCGTGGCTCAGACTCAGTCTGCTGATTG	1135
Qy	1121	GAATCTCTACAGGAATGCATGTTGTGCTCACTACAACAACGCAACAGCATCA	1180
Db	1136	GATCTCTCGCAAGATGCAANTGCTGTGCCACACTACAACTGAAACGCGAAACAGCATAA	1195
Qy	1181	TATATCGATTGAGGGACGGGCTCAGTGCAAGTCTGTGACAGCAACGGCAACAGAGTGT	1240
Db	1196	TATACGATTTGAATGGAAGGGCAATTGATACAAGTGGTGAATTGCAACGCTGAGAGAGTGT	1255
Qy	1241	ACGACGAGGAGCTTCAAGAGGGTCAGTGCTTGTGTGCCACAGAACTTCGCGCTGCTG	1300
Db	1256	TTGATGAGAGCTGCAAGAGGGACGGGTGCTGATCGTGCCCAAAACTTTGTGTGTGCTG	1315
Qy	1301	GAAGAATCCCAAGAGCAGAACTTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCACGA	1360
Db	1316	CAAGATCACAGAGTGACAACCTTCAGTATGTGTCAATCAAGACCAATGATACACCCATGA	1375
Qy	1361	TAGCCAACTTCGGCGGTGAATACTCCGTCTAGATTAACCTGCGGAGGAGGTTGGTGC	1420
Db	1376	TCGGCACTCTTGAGGGGCAACTCAATGTTGTAACGCATTTACAGAGGAAGTGAATTCAGC	1435
Qy	1421	ATTTCATATGCGCTCCAAAGGGACGACGAGGAGCTTAAAGAACAAACAACCCCTTCAAGT	1480
Db	1436	ACACTTTCACCTTAAGAAGCCAGAGGCGAGGCGAGATTAAGAACAACAACCCCTTCAAGT	1495
Qy	1481	TCCTCGCTTCACCGCTCTCAGCAGTCTCCGAGGGCTGTGGCTTA	1523

Db 1136 CATTGGTACAGTGGTGAATTCGATGAGAGAGTGTGATGGAGAGCTGCAAGAGG 1195
Qy 1262 GTCACTGCTGTGTGGTCCACAGAACTTCGCGCTGCTGGAAAGTCCCGAGAGCGAGAACT 1321
Db 1196 GAGGGTGTCTGATGTTCCCAAAACCTTCGGGTGGCTGCAGAAATCCCGAGAGGATAACT 1255
Qy 1322 TCGNATAGTGGGCTTCAAGACAGACTCAAGGCCAGCAGTAGCCAGCTCGCGGTGAAA 1381
Db 1256 TTGAGTATGTGTATTCAAGACCAATGATAGACCTCGATCGGAAACCTTCGAGGGGCAA 1315
Qy 1382 ACTCCGTCATAGATAAATCTCGCGAGGAGTGGTTCGAAATTCATATGGCTCCAAAGGG 1441
Db 1316 ACTCATTTGTAACGCAATTCGAGAGGAGTGTGATGAGCAGCTTTTAACTTAAAGAGCC 1375
Qy 1442 AGCAGGCAAGCGAGCTTAAAGAACAAACCCCTTCAAGTTCTTTCGTTCCACCGTCTCAGC 1501
Db 1376 AGCAGGCCAGCGAGTGAAGAACAAACACCTTTTCAGCTTCTTGTTCACCTCAGGAGT 1435
Qy 1502 AGTCTCCGAGGCTGTGGCTTA 1523
Db 1436 CTCAGAGGAGAGCTGTGGCTTA 1457

RESULT 10

US-10-684-651-12
; Sequence 12, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; APPLICANT: GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-684-651-12

Query Match 33.1%; Score 504; DB 4; Length 1458;
Best Local Similarity 62.0%; Pred. No. 1.2e-140;
Matches 943; Conservative 0; Mismatches 450; Indels 129; Gaps 4;
Qy 5 AGCAACCGGAGGAGAAAGCGGTGCCAGTTCAGCGCTCAATGCGCAGAGACCTTGACAATC 64
Db 62 AGCAGGCAAGCAAAATGATGTCAGATCCAAAAGTGAATGATGCTCCAAACCGGATACCC 121
Qy 65 GCATTGAATCAGAGGCGGTTCATTTGAGACTTTGGAAACCCCAACAAACAGAGAGTTCGAAT 124
Db 122 GTATAGATCGGAAGGTGGTTCATTGAGACATGGAACCTTAACAACAGCCATTCAGT 181
Qy 125 GCGCCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCGCTTCGTAGGCTTCT 184
Db 182 GTGCGGGTGTGCGCTCTCTCGCTGACCCCTTAACCGCAATGCCCTTCGTAGAGCTTCT 241
Qy 185 ACTCCAAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTGATAT 244
Db 242 ACACCAACCGTCCCAAGGAATCTACATACACAGGTAAATGATTTTGGCATGATAT 301
Qy 245 TCCCTGGTTGCTCTAGACACTATGAAGAGCTTCACACAAAGGTCTCGATCTCAGTCCC 304
Db 302 TCCCGGGTGTCTAGCACTTATCAAGAGCGCA----- 335
Qy 305 AAAGACCAACAGAGCGTCTCCAAGGAGAGAACCAAGCCCAACAGCAACGAGATAGTCACC 364
Db 336 -----AGAACTCTCAGCAACGAGGACGAAGCGCCAGAGCGCCCAAGACCGTCACC 382
Qy 365 AGAAGGTGCACCGTTCGATGAGGGTGATCTCATTCAGTTCCCGCTCCACCGGTGTTCTTCT 424
Db 383 AAAGGTACATCGCTTCAGAGAGGGTGATTTGATCGAGTGCCTACTGTGTTGTCATGGT 442
Qy 425 GGCTCTAACACGACCAACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAACAACA 484
Db 443 GGATGTACAAACATGAAGACACTCTCTGTTGTCGCTTCTATTATTGACACCAACAGCT 502
Qy 485 ACAGCAACAGCTTGATGATGTTCCCGAGAGATTCAATTTGGTGGGAACACCGAGCAAG 544
Db 503 TGGAGAACCAAGCTCGACCAAGTCTCTAGGAGATTCTATCTGCTGGGAACCAAGAGCAAG 562
Qy 545 AGTTCTTAAAGGTACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC 604
Db 563 AGTTCTTAAATATCAGCAGCAG----- 585
Qy 605 CATAAGCCCGCAAGTTCAGCTTAGCAAGAAGAGCGTGAAATTTAGCCCTCGAGGACAGC 664
Db 586 -----CAGCAAGGAGGTT 598
Qy 665 ACAGCCGACAGAGAACGAGCAGGACGAAGAAGAAACGAAGTGAAGACATCTTCAGCG 724
Db 599 CCCAAAGCCAGAAAGGAAAGCAACAAAGAAAGAAACGAAGGAAGCAACATATTGAGTG 658
Qy 725 GCTTCACGCGGAGTTCTTGGAAACAGCTTCCAGGTTCAGCAGACAGACAGATAGTCAAA 784
Db 659 GCTTCGCCCTGAAATCTTGAAGAAGAGCGTTC---GGGTGAACATGACAGATAGTAGAA 715
Qy 785 ACCTAAGAGCGGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGGAGGCC 844
Db 716 ACCTAAGAGTGAAGAACGAAGAGGAGGATAGTGGAGCCATTGTGACAGTGAAGGAGGTC 775
Qy 845 TCAGAACTTTGAGCCAG---ATAGAAGAGAGCGTCCCGACGAGAGAGAGGAAATACGATG 901
Db 776 TAAGAGTCAAGCTCCAGCCATGAGGAAGCCACAGCAAGAAGAAGATGATGATGATGAGG 835
Qy 902 AAGATGAATATGAATACGATGAAGAGGATAGAGGGGTGGCAGGGGAAGCAGAGGAGGG 961
Db 836 AAGAGCAGCCACAGTCTCGTGGAGACAGACAAAGTTGCCAACGCCAAAGCAAGAGAGCA 895
Qy 962 GGAATGGTATTGAAGAGACGATCTGCACCGCAGGTCTTAAAGAGAACTTTGGTAGAAACA 1021
Db 896 GAAATGGCAATTGATGAGACCAATTGACCAATGAGACTTCGCCCAAAACATTTGGTCAGAA 955
Qy 1022 GATCCCTTGACATCTTACAAACCCCTCAAGCTGGTTCTACTCAAAACTGCCAACGATCTCAACC 1081

```

Db 956 CATCACTGACATCTAACCCCTCAAGCTGGTAGCATCAACCGCCACCAGCCTTGACT 1015
Qy 1082 TTCTAATACCTAGGTGGCTGGACCTAGTCTGAATATGAAATCTCTACAGGAATGCAT 1141
Db 1016 TCCAGCCCTCTGGCTTCTCAAACTCAGTGGCCAGTATGATCACTCCGCAAGAATGCTA 1075
Qy 1142 TGTGTTGTCGCTCACTCAACACCAAGCGCACACAGCATCATATATCGAATTGAGGGGACGGG 1201
Db 1076 TGTTCGTGCCACACTACACCTGAAACGCGAACAGCATATATACGCAATTGAATGGCGGG 1135
Qy 1202 CTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGG 1261
Db 1136 CATTTGTCACAAGTGGTGAATTGCAATGCTGAGAGAGTGTGATGAGAGCTGCAAGAGG 1195
Qy 1262 GTCACGTGCTTGTGGTGCCACAGAACTTCGCCGTCTGCTGGAAGTCCAGAGCGAGAACT 1321
Db 1196 GAGGGGTGCTGATCGTTTCCACAAAACCTTTGCGGTGGCTGCAAAATCCCAGAGCGATACT 1255
Qy 1322 TCGAATACGTGGCAATCAAGACAGACTCAAGGCCCAGCATAGCAACCTCGCCGGTGAAA 1381
Db 1256 TTGAGTATGTCAATCAAGACCAATGATAGACCTCGATCGGAACCTTGCAGGGGCAA 1315
Qy 1382 ACTCCGTATAGATAACCTGCGGAGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGGG 1441
Db 1316 ACTCATTTGTTGAACGCAATTTGCCAGAGGAAGTGATTCAGCACACTTTTAACTTAAAGAGCC 1375
Qy 1442 AGCAGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGC 1501
Db 1376 AGCAGCCAGGAGGTGAAGAACAAACCCCTTCAGCTTCTTGTTCACCTCAGGAGT 1435
Qy 1502 AGTCTCCGAGGCTGTGGCTTA 1523
Db 1436 CTCAGAGGAGAGCTGTGGCTTA 1457
```

RESULT 11

```

US-09-108-010B-15
; Sequence 15, Application US/09108010B
; Patent No. 6362399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
```

```

; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-108-010B-15

Query Match      13.0%; Score 198; DB 3; Length 1551;
Best Local Similarity 51.4%; Pred.No.1.le-48;
Matches 699; Conservative 0; Mismatches 60; Indels 61; Gaps 8;

Qy 14 AGGAGAACGCGTCCAGTTCCAGCGCCTCAATGCCAGAGACCTGACAAATCCGATTGAAT 73
Db 86 AGTTCAACGAGTGCCTCACTCAACACCTCAAGCGCTTGAACCCGACCCACCGCTTGAT 145
Qy 74 CAGAGGGCGGTTTACATTGAGACTTGGAAACCCCAACACAGGAGTTTGAATGCGCGGG 133
Db 146 CGGAAGTGGTCTTATTGAAACATGGAACCTCTCAACACCCCTGAGCTGCAATGCGCGGTG 205
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAAGCCCTTCGTAGGCTTCTTACTCCAATG 193
Db 206 TCACTGTTTCCAAACGCAACCTCAACCGCAAGCGCTCCCACTTGCCATCTTACTTACCTT 265
Qy 194 CTCGCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTGATATTCCCTGGTT 253
Db 266 ATCCCAATGATCATTTGCTGTTCAAGGAAGGGAGCAATTTGATTTGCAATTTCCGGAT 325
Qy 254 GTCTTAGACATATGAAGAGCTCTACACAAGGTCGTGATCTCAGTCCCCAAGACCAAC 313
Db 326 GTCCGAGACGTTTGAGAAGCCACA-----ACAAAC 355
Qy 314 CAAGAGCTCTCAAGAGAGAGACCAAGCAACAGCAGATAGTACACAGAAGGTGC 373
Db 356 AATCAAGCAGAAGAGGCTCAAGGTCAAGCAGCACTAACAAGACAGTACCAGAAGATTC 415
Qy 374 ACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCCCAACCGGTGTTGCTTTCTGGCTCTACA 433
Db 416 GTCACTTCAATGAAGAGAGACGTACTAGTATCTCTTGGTGTCTTCTTACTGACCTATA 475
Qy 434 AGCAGCAGACACTGATGTTGTTGTTTCTTCTTACTGACACCAACAAACAGCAACCC 493
Db 476 ACACCTGGCGATGAACCAAGTGTGTGCATCAGTCTCTTGACACCTCCCACTTCAACAATC 535
Qy 494 AGCTTGATCAGTTCCCGCAGGATTTCAATTTGCTGGGAACACGCGAGCAAGAGTTCTTAA 553
Db 536 AGCTTGATCAAAACCCCGAGATTTTACCTTGTGGGAACCCAGATATAGAGCATCCCG 595
Qy 554 GGTACCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACAGCC 613
Db 596 AGACCATGCAACA--CAGCAGCAGCAGAGAGTATGTTGGACGCAAGCGGGGCAACA 653
Qy 614 CGCAAGTCAGCTTACAGCAAGAG--AGCGTGAATTTAGCCCTCGAGGACAGCAGCCCG 672
Db 654 CGCAGCAGCAGGAGGAAGGTGGCAGTGTCTCAGTGGCTTTCAGCAACATTTCTTAC 713
Qy 673 AGAGAACGAGCAGGACAGAAAGAAACGAAGGTGGAAACATCTTCAG----CGGCTT 728
Db 714 ACAATCTTCAACACCAACGAGGACACAGTGTAGAAACTTCGGTCTCCAGATGACGAAG 773
Qy 729 CACGCCGAGTTCTCGGAACAGCCCTTCAGGTTGACGACAGACAGATAGTGCAAAACCT 788
Db 774 GAAGCAGATCGTGACAGTGGAGGGGCGCTCAGCGTTATCAGCCCAAGTGCAGAAACA 833
Qy 789 AAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTGACAGTGGAGGGAGGCTCAG 848
Db 834 AGAAGACGAAGACGAAGACGAAGAGCAAGATATGAGCGGAGCGCCCTCTTATCTCCACG 893
Qy 849 AATCTTTGAGCCCGCAGATAGAAAGAGAGCGTGC---GACGAAGAGAGGAATACGATGAAGA 905
Db 894 ACACCAAGCCATGGAAAGCATGAAGATGACAGGACGAGGACGAAGAGAGATCAACCC 953
```


QY 906 TGAATATGAATACATGAGAGAGTAGAAG-----GGTGGCAGGGG 947
Db 954 TCCTCTGATACACCTCCACAGGACCAAGCAGGCCCGGAAACAAGAACCACTGGAG 1013
QY 948 AAGCAGAGGAGGGGAATGATTTGAAGAGAGATCTGACACCGCAAGTGTCTAAAGAA 1007
Db 1014 AGATGTCAGACTAGAAATGGGTTGAGGAATATTTGACCATGAGCTTCACAGAA 1073
QY 1008 CATTGGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGC 1067
Db 1074 CATTGCTCGCCCTTCACGCTGACTTCTACAAACCCAAAGCTGGTGCATTAGCACCT 1133
QY 1068 CAAGGATCTCAACTTCTAATAGTGGCTTGGACCTAGTCTGATGAAATCT 1127
Db 1134 CAACAGTCTCACCTCCACCCCTCGCCCAATTGCGACTCAGTGGCCCAATATGTTGCT 1193
QY 1128 CTACAGGAATGCAATGTTTCTGCTCACTACAAACCCAAAGCTGGTGCATATATCG 1187
Db 1194 CTACAGGAATGGAATTTACTCTCAGATTTGGAATCTTGAACCGCAAGTGTGAG-ATGA 1252
QY 1188 ATTGAGGGGAGGCTCACGTGCAAGTCTGAGACAGCAAGCAAGCAAGTGTACGACGA 1247
Db 1253 CTCGAGGGAAGG--AAGAGTAGAGTGTGAATCTGCAAGGGAATGCAAGTGTGACGG 1310
QY 1248 GAGCTTCAAGAGGTCACGTCTTGTGTTGTTGTCACAGAACTTCGCGTCTGCTGAAAGTC 1307
Db 1311 TGAGCTAAGGAGGGGCAATTTGCTAGTGTGTCGCGGCAAGCCCGCGTGTGCTGAGCAAG 1370
QY 1308 CCAGAGCGAGAACTTCAATACGTGGCATTCAGACAGAC 1347
Db 1371 GGGAGAACAGGATTTGAATATGATGTTTCAAGACACAC 1410

RESULT 12

US-09-758-652-15
; Sequence 15, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-758-652-15

Query Match
Best Local Similarity 51.4%; Pred. No. 1.1e-48;
Matches 699; Conservative 0; Mismatches 600; Indels 61; Gaps 8;

QY 14 AGGAGAACCGGTGCGAGTTCCAGCGCTCAATGCGCAGAGACCTGACAAATCGCATTTGAAT 73
Db 86 AGTTCAACGAGTGCCTCAACCACTCAACGCTTGAACCCGACACCGGTTGAGT 145
QY 74 CAGAGGGCGGTTACATTTGAGACTTGGAAACCCCAACCAACAGGAGTTTCGAATGCGCGGCG 133
Db 146 CCGAGGTGGTCTTATTGAAACATGGAATCTCAACACCTCTGAGCTGCAATGCGCGGTG 205
QY 134 TCGCCCTCTCTCGCTTAGTCTTCGCGCGCAACGCCCTTCGTAGGCTTTCTACTCCAATG 193
Db 206 TCACTGTTTCCAAACGACCTCAACCGCAAGGCTCCCACTTGCATCTTACTTACCTT 265
QY 194 CTCGCCAGGAGATCTTCATCCAGCAGGAAGGAGTACTTTGGTTGATATTCCTGTT 253
Db 266 ATCCCAAAATGATCATTTGCTTCAAGGGAAGGAGCAATTTGATTTGCAATTTCCGGAT 325
QY 254 GTCCTAGACACTATGAAGAGCCTCAACACAAAGTCTGCTGATCTCAGTCCCAAGAGCAC 313
Db 326 GTCGCCAGGAGTTTGAAGCCACA-----ACAAAC 355
QY 314 CAAAGAGTCTCAAGGAGGAAGCAAGCAACAGCAAGAGATGTCACCAAGAGGTGC 373
Db 356 AATCAAGCAGAAAGGCTCAAGTCAAGCAGCAAGCACTACAGACAGTCAACAGAGATTC 415
QY 374 ACGTTTCATGAGGTGATCTCATTTGAGTTTCCCAAGGTTGCTTCTGCTCTACA 433
Db 416 GTCATCTCAATGAAGGAGAGCTACTAGTATCTCTTGGTTTCTTACTTGACCTATA 475
QY 434 ACGACCAACGACACTGATGTTGTTGCTTCTTCTTACTGACACCAACAAACAGCAAC 493
Db 476 AACTGGCATGAAACGATTTGCTTCCATCAGTCTCTTGCACCTCCAACTTCAACATC 535
QY 494 AGCTTGATCAGTTTCCCGAGGATTCATTTGGCTGGGAACACGAGAGAGTCTTTAA 553
Db 536 AGCTTGATCAAAACCCCGAGATTTTACTTCTGCTGGAAACCCAGATATAGAGCATCCCG 595
QY 554 GGTACAGCAACAAAGCAGACAAAGCAGAGAGAGCTTACCATATAGCCCATACAGCC 613
Db 596 AGACCATGCAACAA--CAGCAGCAGCAGAAAGTCTATGTTGGACGCAAGCGGGCAACA 653
QY 614 CGCAAGTCAGCTAGACAAAGAG--AGCGTGAATTTAGCCCTCGAGGACAGCAGCGCG 672
Db 654 CGACAGCAGGAGGAAGAGGTGGCAGTGTCTAGTGGCTTCAGCAACATTTCTTAGC 713
QY 673 AGAGAACGAGCAGGGAACAAGAAAGAAACAGAGGTGGAAACATCTTCAG-----CGGCTT 728
Db 714 AATCTCTTCAACACCAACGAGGACAGCTGAGAAACTTCGGTCTCCAGATCAGAAAG 773
QY 729 CAGCCGGAGTCTCTGGAAACAGCTTCCAGGTTGACGAGCAGACAGATAGTGAACACCT 788
Db 774 GAAGCAGATCGTGACAGTGGAGGGGCGCTCAGCGTTATCAGCCCAAGTGGCAAGACA 833
QY 789 AAGAGCGGAGACCGAGAGTGAAGACAGGAGGAGCATTGTGACAGTGAAGGAGGCGCTCAG 848
Db 834 AGAAGCAGAGAGCAAGACAGCAAGCAAGAAATATGGAACGAGCGCTCTTATCTCTCAGC 893
QY 849 AATCTTTGAGCCCGCAGATAGAAAGAGAGCTGCC---GACGAAGAGAGAGAAATACGATGA 905
Db 894 ACGACCAAGCCATGGAAGAGCATGAAGATGACGAGGAGCAGGACGAAGAGAGATCAACC 953
QY 906 TGAATATGAATACGATGAAGAGATAGAAG-----CGGTGGCAGGGG 947

Db 954 TCCTGCTGATCACCTCCACAGCGACCAGCGGCGGACAAACAAAGAACCCAGCTGGAAG 1013
Qy 948 AACGAGGCGAGGGGNAATGGTATTGAAGAGAGCATCTGCACCGCAAGTGTCTAAAGAA 1007
Db 1014 AGGATGTCAGACTAGAAATGGGGTTGAGGAAATATTTGCACCATGAAGCTTCACGAGAA 1073
Qy 1008 CATTGGTAGAAACAGATCCCTCAGCATCTACAACTCTCAAGCTGGTTCACTCAAAACTGC 1067
Db 1074 CATTGCTGCCCTTCAGTCTGACTTCTACAACTCCAAAGCTGGTCGATAGCACTT 1133
Qy 1068 CAACGATCTCAACTCTTAATCTTAGTGGCTTGGACCTAGTGTGGAATATGGAATCT 1127
Db 1134 CAACAGTCTCACCTCCAGCCCTCGCCCAATTCGGACTCAGTGCCTCAATATGTTGCTCT 1193
Qy 1128 CTACAGGATGCAATGTTTGTCTCTACTACACCAACGACACAGCATCATATATCG 1187
Db 1194 CTACAGGAATGGAATTTACTCTCCAGATTGGAATCTTGAACGCGAAACAGTGTGACG-ATGA 1252
Qy 1188 ATTGAGGGGACGGGCTCAGTGCAGTCTGAGACAGCAACGCGCAACAGAGTGTACGACGA 1247
Db 1253 CTCGAGGGAAGG--AAGAGTTAGAGTGGTGAATCTGCCAAGGAATGCAGTGTTCGACGG 1310
Qy 1248 GGAGCTTCAAGAGGGTCACGTGTTGTGTGCCACAGAACTTCGCCGTGCTGGAAGTTC 1307
Db 1311 TGAGCTAAGGAGGGGACAAATTGCTAGTGGTCCGCGAGAACCCCGCGGTGGCTGAGCAAGG 1370
Qy 1308 CCAGAGCGAGAACTTCGATACGTGCGCATTCAGACAGAC 1347
Db 1371 GGGAGAACAGGATGGAATATGTTAGTGTTCAGACACAC 1410

RESULT 13

US-10-684-651-15
; Sequence 15, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; APPLICANT: GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
;

; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-684-651-15
Query Match 13.0%; Score 198; DB 4; Length 1551;
Best Local Similarity 51.4%; Pred. No. 1.1e-48;
Matches 699; Conservative 0; Mismatches 600; Indels 61; Gaps 8;
Qy 14 AGGAGAACCGTGCAGCTTCCAGCGCTCAATGCGCAGAGACCTGCACATCGCATTTGAAT 73
Db 86 AGTTCAACGAGTGCCAACTCAACAACCTCAACGCGTTGGAACCCGACACCGCGTTGAGT 145
Qy 74 CAGAGGGGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTTGAATGCGCGGGCG 133
Db 146 CGAAAGGTGCTTATTGAAACATGGAATCTTCAACACCTGAGCTGCAATGCGCGGTG 205
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCTTCTGTAAGGCTTTCTACTCAATG 193
Db 206 TCACCTGTTTCCAAACGACCTTCAACCGCAACGGCTCCCACTTGCATCTTACTTACCTT 265
Qy 194 CTCCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTGATATTTCCCTGTT 253
Db 266 ATCCCCAAATGATCATTTGCTGTTCAAGGGAAGGAGCAATTGGATTTCATTTCCGGGAT 325
Qy 254 GTCTTAGACATATGAAGAGCTCTACACAAAGGTCGTGATCTCAGTCCCCAAAGACCAAC 313
Db 326 GTCCCGAGACGTTTGAGAAGCCACA-----ACAAAC 355
Qy 314 CAAGACGCTCCCAAGGAGAGACCAAGCCAAACAGACAGATAGTACCAGAAAGGTGC 373
Db 356 AATCAAGCAGAAGAGGCTCAAGGTCAAGCAGCACTACAAGACAGTACCAGAAAGATTC 415
Qy 374 ACCGTTTCGATGAGGGTGTATCTCATTTGCAGTTTCCCAACGGGTGTTGCTTTCTGGCTCTACA 433
Db 416 GTCACTTCAATGAAGAGAGACGTTACTAGTATCTCTTGGTGTCTTACTTGACCTATA 475
Qy 434 ACGACAGACACTGATGTTGTTGCTGTTTCTTCTACTGACACCAACAAACAGCAACCC 493
Db 476 ACACCTGGCGATGAACCACTGTTTGCATCAGTCTCTTTGACACCTTCAACTTTCAACAATC 535
Qy 494 AGCTTTGATCAGTTCCCGCAGGAGATTCATTTTGGCTGGGAAACACGCGAGCAAGAGTTCTTAA 553
Db 536 AGCTTGATCAAAACCCAGAGTATTTTACCTTGTCTGGGAAACCCAGATATAGAGCATCCCG 595
Qy 554 GGTACAGCAACAAAGCAGACAAAGCAGACGAAAGCTTTACCATATAGCCCATACAGCC 613
Db 596 AGACCATGCAACA--CAGCAGCAGCAGAGAGTCTATGTTGGACGCAAGCAGGGGCAACA 653
Qy 614 CGCAAGTCAGCTAGACAAAG--AGCGTGAATTTAGCCCTCGAGGACAGACAGCCGC 672
Db 654 CGCAGCAGCAGGAGGAAAGGTGGCAGTGTGCTCAGTGGCTTCAGCAACATTTCTTAGC 713
Qy 673 AGAGAACGAGCAGGACAAAGAAAGAAACGAAAGGTGGAAACATCTTCAG-----CGGCTT 728
Db 714 ACATCTCTTCAACACCAAGCAGGACACAGCTGAGAACTTCGGTCTCCAGATGACGAAG 773
Qy 729 CACGCCGGAGTTTCTTGGAAACAGCCCTTCAGGTTTGACGACAGACAGATAGTGCAAAACCT 788
Db 774 GAAGCAGATCGTGACAGTGGAGGGGCTCAGCGTTATCAGCCCAAGTGCAGAAACA 833
Qy 789 AAGAGCGAGACCGAGAGTGAAGAGAGGAGCCATTTGTGACAGTGGGGAGGCTCAG 848
Db 834 AGAAGACGAAGCAAGAGACGAGAGCAAGAAATATGAGCGGACGCCCTCTTTATCTCCACG 893
Qy 849 AATCTTTGAGCCAGATAGAAAGAGACGCTGCC---CACGAAGAGAGGAAATACGATGAAGA 905
Db 894 ACGACCAAGCCATGGAAGCATGAAGATGAACGAGGACGAGGACGAGAAAGAGATCAACC 953
Qy 906 TGAATATGAATACGATGAAGAGGATAGAAG-----GCGTGGCAGGGG 947

Db 954 TCCTCTGATCACCCTTCCACAGCCAGCAGCGCCGGAACCAACAGAACACACGTGGAAG 1013
Qy 948 AAGCAGAGCGAGGGGAATGGTATTGGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAA 1007
Db 1014 AGGATGTCAGACTAGAAATGGGTTGAGGAAATATTTGACACCATGAAGCTTCACGAGAA 1073
Qy 1008 CATTGGTAGAACAAGATCCCTGACATCTCAACACCTCAAGCTGTTCACTCAAAATGC 1067
Db 1074 CATTGCTCGCCCTTCACGTGCTGACTTCTCAACACCAAAAGCTGGTCGCAATTAGCACCCCT 1133
Qy 1068 CAACGATCTCAACCTTCTAATACCTAGTGGCTTGGACCTAGTCTGAATATGAAATCT 1127
Db 1134 CAACAGTCTACCTCCAGCCCTCCGCAATTCGGAATTCGGAATTCGGAATTCGGAATTCG 1193
Qy 1128 CTACAGGAATGCAATGTTTTCGCTCACTCAACACCAACGCAACAGAGTGTACGATCAT 1187
Db 1194 CTACAGGAATGGAATTTACTCTCCAGATTGGAATTCGAACGCAACAGTGTGACG-ATGA 1252
Qy 1188 ATTGAGGGAGCGGCTCACGTGCAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGT 1247
Db 1253 CTCGAGGAAAGG--AAGAGTTAGAGTGGTGAATCTGCAAGGGAATCGAGTGTTCGACGG 1310
Qy 1248 GGAGCTCAAGAGAGGTCAAGTGTGTTGGTCCACAGAACTTCGCCGTTCGCTGGAAGTC 1307
Db 1311 TGAGCTAAGGAGGGAACAATGCTAGTGGTCCGCAAGACCCCGCGTGGCTGAGCAAGG 1370
Qy 1308 CCAGAGCGAGAACTTCGAATACGTGGCAATTCAGACAGAC 1347
Db 1371 GGGAGAACAGGATTGGAATATGATGTTTCAAGACACAC 1410

RESULT 14

US-09-108-010B-14

; Sequence 14, Application US/09108010B

; Patent No. 6362399

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WINDOWS 95

; SOFTWARE: MICROSOFT WORD 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/108.010B

; FILING DATE: 30-Jun-1998

; CLASSIFICATION: <unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/019,940

; FILING DATE: JUNE 14,1996

; ATTORNEY/AGENT INFORMATION:

; NAME: LYNNE M. CHRISTENBURY

; REGISTRATION NUMBER: 30,971

; REFERENCE/DOCKET NUMBER: BB-1071-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-5481

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1689 base pairs

; TYPE: nucleic acid

;

;

;

;

;

;

;

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-108-010B-14

Query Match 10.1%; Score 153.2; DB 3; Length 1689;

Best Local Similarity 49.7%; Pred. No. 3.2e-35; Indels 156; Gaps 7;
Matches 726; Conservative 0; Mismatches 578;

Qy 14 AGGAGAACGGTCCCGCTTCCAGCGCTCAATCGCAGAGACCTGACAATTCGATTTGAAT 73
Db 83 AGCTCAACGAGTCCCACTCAACAACTCAACGGTTGGAACCCGACCCGCGTTGAGT 142
Qy 74 CAGAGGGCGTTACATTTAGACTTTGAAACCCCAACACAGGAGTTTGAATGCGCGCG 133
Db 143 TCGAAGTGGTTTGAATTTCAACATGGAACCTCTCAACACCTGAGCTGAAATGCGCGGTG 202
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCTAGTGGCTTTCTACTCCAATG 193
Db 203 TCACTGTTTCCAAACTCACCTCAACCGCAATGGCTTCCACTTGCATCTTACTCACCTT 262
Qy 194 CTCCCAGAGAGACTTTTCATCCAGCAAGGAAGGGGATACTTTGGGTTGATATTCCTCGT 253
Db 263 ATCCCGGATGATCATATCGCCCAAGGGAAGAGCACTGCAGTGCA---AGCCAGGAT 319
Qy 254 GTCTTAGACACTATAGAGAGCTTCAACACAAAGGTCTGATCTCAGTCCCAAGACCAC 313
Db 320 GTCTGAGAGCTTTGAGAGGCCACAAGAACAA-----TCAACAGAAAGAGCTC 368
Qy 314 CAAGAGCTCTCAAGGAGAGAACCAAGCAACAGCAAGAGATAGTACCAAGAGGTGC 373
Db 369 AAGGTCTG-----CAGAAGCAGCAGCTACAGGACAGTCCACAGAAGATTC 412
Qy 374 ACCGTTTCGATGAGGTGATCTCATTGCACTTCCACCGGTGCTTCTTCTGCTCTACA 433
Db 413 GTCATCTCAATGAGAGAGAGCTACTCGTATCTCTCTGCTGTTCTTACTGACCTATA 472
Qy 434 ACACCAACGACACTGATGTTGTTGCTTTCTTTACTAGACACCAACAAACAGCAAC 493
Db 473 ACCTGGCGAGAACCACTGTTGTTGCTTCTTCTGACACTCTTAACCTCAATAAACC 532
Qy 494 AGCTTGTAGTTCCTCCAGAGATTTCAATTTGGTGGACACGAGAGAGTCTTTAA 553
Db 533 AGCTTGTATCAACCCCTTAGGGTATTTTACCTTCTGCGAACCCAGATATAGAGTACCAG 592
Qy 554 GTTACCAAGCAACAAAGCAGACA----- 575
Db 593 AGACCATGCACAAACAAACAGCAGCAAGAAAGTCAATGGTGGACGACAGCGGGGCACACC 652
Qy 576 -----AAGCAGACGAAGAGCTTACATATAGCCCATACAGCCCGCAAGTCAAGCTAG 629
Db 653 AGCAGGAGGAAGAGGAAGAGGTGGCAGCGTCTCAGTGGCTTTCAGCAACACTTCTTGG 712
Qy 630 ACAAGAGAGCGTGAATTTAGCCCTCGAGGACAGCAGCGCGCAGAGAAACATTCAGTCT 689
Db 713 CACAATCTTTCACACCAACAGAGACATAGCTGAGAAATCTTCCAGACGACGAAA 772
Qy 690 AGAAGAGAAACGAAGGTGGAACATCT--TCAGCGCTTTCAGCGCGAGTCTCTGGAAC 748
Db 773 GGAAGCAGATCGTGACAGTGGAAAGAGGTCTCAGCGTTATCAGCCCTCCAGTGGCAAGAC 832
Qy 749 AAGCCTTTCAGGTTGACGACAGACAGATAGTGC AAA----- 785
Db 833 AACAGATGAAGATGAAGATGAAGACGAAGATGATGAAGATGAACAAATTCCTCTCAC 892
Qy 786 -----CCTAAGAGCGGAGACCGAGTGAGAGAGGAGCCATTGTGACAGTAGGG 838
Db 893 CTCTCGCCGACCAAGCCATGGAAAGCGTGAAACAAGACGAGGACGAGGACGAAGATGAAG 952
Qy 839 GAGGCTCAGAATCTTGACCCAGATAGAAAGAGAGTGTCCGACGAGAGAGGAAATACG 898
Db 953 ATAAACCTCTGCTCTAGTCTGACCAAGCAAGAAAGCGTGAAACAGACAGGACGAGGACG 1012

QY 899 ATGAAGATGAATATGATACGATCAAGAGGATAGAAAGCGGTGCGAGG----- 945
Db 1013 AGGACGAAGATGAAGATGAAGATCAACCTCGCAGAGCGCGAATGGAGATCGAAAAGA 1072
QY 946 -----GGAACGAGGCGAGGGGAATG 967
Db 1073 CACAACCCAGAAGACCTAGACAAGAAGAACCCAGTGAAGAGGATGCGAGACAAGAAACG 1132
QY 968 GTATTGAAGAGAGATCTGCACCGCAAGTGTCTAAAGAGACATTTGTAGAAAACAGATCCC 1027
Db 1133 GGGTTGAGGAAATATCTGCACCTTGAAGCTTCACGAGAACATTTGTCGCCCTTCACGG 1192
QY 1028 CTGACATCTACAAACCTCAAGCTGGTTCACTCAAAAACCTGCCAACGATCTCAACCTTCTAA 1087
Db 1193 CTGACTCTACAAACCTTAAAGCTGGTTCGCAATTAGTACCTTCAACAGCTCACCTCCAG 1252
QY 1088 TACTTAGTGGCTTGGACCTAGTGTGTAATATGAAATCTCTACAGGAATGCATTTTGG 1147
Db 1253 CCTCGCCCAATTCCAACTCAGTGCCTCAATATGTTGCTCTACAAGAAATGGAATTTACT 1312
QY 1148 TCCTCACTACACCAACGACACAGCATCATATATCGATTGAGGGGACGGCTCAG 1207
Db 1313 CTCACATATGGAATCTGAATGCAACAGTGTGATCTATGTGACTCGAGGACAAAGGAAG 1372
QY 1208 TGCAAGTCTGTGAGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAG 1267
Db 1373 TTAGAGTTGTAAGTCTGCAAGGAATGCAGTGTTCGACGGTGAAGTTCAGGAGGGAAT 1432
QY 1268 TGCTTGTGGTCCACAGAACTTCGCGTGTGCTGGAAGTCCAGACGAGAACTTCGAAT 1327
Db 1433 TGCTGTGTGTACACAGAACTTCGTGTGGCGAGCAAGCGGAGAAACAGGATTCGAAT 1492
QY 1328 ACGTGGCATTCAGACAGAC 1347
Db 1493 ACATAGTATTCAAGACAC 1512

RESULT 15

US-09-758-652-14
; Sequence 14, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENSEN
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-758-652-14

Query Match 10.1%; Score 153.2; DB 4; Length 1689;
Best Local Similarity 49.7%; Pred. No. 3.2e-35;
Matches 726; Conservative 0; Mismatches 578; Indels 156; Gaps 7;

QY 14 AGAGAAACGCGTTCAGCGCTCAATGCGCAGAGACCTGACATCGCATTTGAAT 73
Db 83 AGCTCAACGAGTGCACAACTCAACCGCTTGGAAACCCGACCCACCGCTTGA 142
QY 74 CAGAGGGCGTTTACATTGAGACTTGGAAACCCCAACACGAGAGTTTCGAATGCGCGGCG 133
Db 143 TCGAAGGTGGTTTGAATTAACATGAACTCTCAACACCTGAGCTGAATGCGCGGCG 202
QY 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCTGTAGGCTTTTCTACTCCAATG 193
Db 203 TCACTGTTTCCAAACTCACCCCTCAACCGCAATGGCTCCACTTGCATCTTACTCACCTT 262
QY 194 CTCGCCAGAGATCTTCATCCAGCAAGGAGGGGATACCTTTGGGTGTGATTTCCCTGGTT 253
Db 263 ATCCCGGATGATCATCGCCCAAGGGAAGAGGACACTGCGAGTGCA---AGCCAGGAT 319
QY 254 GTCCTTAGACACTATGAAGAGCTCAACACAGGTGCTCGATCTCAGTCCCAAGAACAC 313
Db 320 GTCCTGAGACGTTTGAAGGCCACAGAAACA-----TCACAGAAAGGGCTC 368
QY 314 CAAGACGTCTCAAGGAGAAGACCAAGCCAAACAGACAGATAGTCCACAGAGGTGC 373
Db 369 AAGGTGCG-----CAGAAGCAGCAGCTACAGGACAGTACCAGAAAGATT 412
QY 374 ACGTTTCGATGAGGTGATCTCATTCGATGCGATTCACCGGTGTTGTTCTCGGCTCTACA 433
Db 413 GTCACCTTCAATGAAGGAGACGTAACCTGTTCTCTGTTGTTCTTACTGGACCTATA 472
QY 434 AGCACACGACACTGATGTTGTTGCTTCTTCTTACTGACACCAACCAACACCAAC 493
Db 473 ACACCTGGCGATGAACCGAGTTGTTGCCATCAGTCTTCTTGACACCTTAACTCAATAAC 532
QY 494 AGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAGCAAGAGTTCTTAA 553
Db 533 AGCTTGATCAAAACCCCTAGGGTATTTTACCTTGTCTGGGAACCCAGATATAGAGTACCCAG 592
QY 554 GGTACAGCAACAAAGCAGACA----- 575
Db 593 AGACCATGCAACAACAACACAGCAGAAAAGTCAATGGTGGACGCAAGCGGGCAACACC 652
QY 576 -----AAGCAGACGAAGAAGCTTACCATATAGCCCATACAGCCCGCAAGTCAGCTAG 629
Db 653 AGCAGAGGAAGGAAGGAAGGTGGCGGCTCTCAGTGGCTTCAGCAAAACACTTCTTGG 712
QY 630 ACAAGAAGAGCGTGAATTTTAGCCCTCGAGGACAGACAGCCCGCAGAGAACGAGCAGGACA 689
Db 713 CACAATCTCTTCAACACCAACGAGGACATAGCTGAGAAAATCTTCAGTCTCCAGACGAGAAA 772
QY 690 AGAAGAAGAAACGAGGTGGAAAACATCT-TCAGCGGCTTCAGCGGGAGTTCTCTGGAC 748
Db 773 CGAAGCAGATCGTGACAGTGGGAAGGAGGTCTCAGCGTTATCAGCCCTAAGTGGCAAGAAC 832
QY 749 AAGCCCTTCAGGTTGACGACACACAGATAGTGCAAAA----- 785
Db 833 ACAAGATGAAGATGAAGATGAAGCAAGATGATGAAGATGAAACAAATTCCTCTCACC 892
QY 786 -----CCTAAGAGCGGAGACCGAGAGTGAAGAAAGAGGGAGGCCATTGTGACAGTGAGGG 838

```
Db      893  CTCCTCCGCGACCAAGCCATGGAAAGCGTGAACAAAGACGAGGACGAGACGAAGATGAAG 952
Qy      839  GAGGCCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTCGCGACGAAGAAGAGGAATACG 898
Db      953  ATAAACCTCGTCCTAGTCGACCAAGCCAAAGGAAAGCGTGAACAAAGACGAGGCCAGGACG 1012
Qy      899  ATGAAGATGAATATGAATACGATGAAGAGGATGAAGGCGTGGCAGG----- 945
Db      1013  AGGACGAAGATGAAGATGAAGATCAACCTCGCAAGAGCGCGCAATGGAGATCGAAAAAGA 1072
Qy      946  -----GGAAAGCAGAGGCGAGGGGGAATG 967
Db      1073  CACAACCCAGAGACCTAGACAAGAAGAACACGCTGAAGAGGATGCGAGACAAGAAGCG 1132
Qy      968  GTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAACATTGGTAGAAACAGATCCC 1027
Db      1133  GGGTTGAGGAAATATCTGCACCTTGAAGCTTACAGAGAACATTGCTCGCCCTTCACGCG 1192
Qy      1028  CTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACTGCCAAGCATCTCAACCTTCTAA 1087
Db      1193  CTGACTTCTACAACCCCTAAAGCTGGTCGCATTAGTACCTCAACAGGCTCACCCCTCCAG 1252
Qy      1088  TACTTAGTGGCTTGGACCTAGTCTGAATATGAAATCTCTACAGGAATGCAATTGTTG 1147
Db      1253  CCTCCGCGAATTCCAACTCAGTGCCCAATATGTTGCTCTACAAGATGGAAATTTACT 1312
Qy      1148  TCGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGGGCTCACG 1207
Db      1313  CTCCACATTTGGAATCTGAATGCAACACAGTGTGTATGTGACTCGAGGACAAAGGAAAGG 1372
Qy      1208  TGCNAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACG 1267
Db      1373  TTAGAGTTGTGAACCTGCCAAGGGAATGCAGTGTTCGACGGGTGAGCTTAGGAGGGGACAAT 1432
Qy      1268  TGCTTGTGGTGCCACAGAACTTCGCCGTCGCTGGAAAGTCCAGAGCGAGAACTTTCGAAT 1327
Db      1433  TGCTGGTGGTACCAAGACAACTTCGTGGTGGGAGCAAGCGCGGAGAACAGGATTCGAAT 1492
Qy      1328  ACGTGGCAATTCAGACAGAC 1347
Db      1493  ACATAGTATTCAAGACACAC 1512
```

Search completed: August 24, 2005, 03:57:19

Job time : 267.731 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 921.248 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 CGGAGCAACCGAGGAGGAGAA.....ctccaggagctgtggttaa 1524

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_5/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1524	100.0	1524	10	US-09-731-375A-3
2	1524	100.0	1524	14	US-10-728-806-5
3	1524	100.0	1524	17	US-10-100-303A-89
4	1524	100.0	1524	20	US-10-728-323-3
5	1524	100.0	1524	20	US-10-728-051-3
6	1524	100.0	1524	21	US-10-899-551-5
7	1391.2	91.3	1853	21	US-10-958-324-4

8	517.4	34.0	1446	9	US-09-758-652-13	Sequence 13, Appl
9	517.4	34.0	1446	18	US-10-684-651-13	Sequence 13, Appl
10	517.4	34.0	1446	19	US-10-757-074-13	Sequence 13, Appl
11	517.4	34.0	1446	19	US-10-757-155-13	Sequence 13, Appl
12	517.4	34.0	1446	19	US-10-757-667-13	Sequence 13, Appl
13	517.4	34.0	1673	18	US-10-425-114-7280	Sequence 7280, Ap
14	517.4	34.0	1689	18	US-10-425-114-8531	Sequence 8531, Ap
15	517.4	34.0	1694	18	US-10-425-114-8592	Sequence 8592, Ap
16	517.4	34.0	1697	18	US-10-425-114-7583	Sequence 7583, Ap
17	517.4	34.0	1766	18	US-10-424-599-28860	Sequence 28860, A
18	515.8	33.8	1671	18	US-10-425-114-7594	Sequence 7594, Ap
19	513	33.7	1431	21	US-10-409-993-2	Sequence 2, Appl
20	513	33.7	1434	21	US-10-409-993-16	Sequence 16, Appl
21	513	33.7	1458	21	US-10-409-993-10	Sequence 10, Appl
22	513	33.7	1488	9	US-09-758-652-11	Sequence 11, Appl
23	513	33.7	1488	18	US-10-684-651-11	Sequence 11, Appl
24	513	33.7	1488	19	US-10-757-074-11	Sequence 11, Appl
25	513	33.7	1488	19	US-10-757-155-11	Sequence 11, Appl
26	513	33.7	1488	19	US-10-757-667-11	Sequence 11, Appl
27	513	33.7	1694	18	US-10-425-114-7674	Sequence 7674, Ap
28	513	33.7	1716	18	US-10-425-114-8556	Sequence 8556, Ap
29	513	33.7	1718	18	US-10-425-114-7642	Sequence 7642, Ap
30	513	33.7	1721	18	US-10-425-114-7670	Sequence 7670, Ap
31	513	33.7	1724	18	US-10-425-114-7613	Sequence 7613, Ap
32	513	33.7	1724	18	US-10-425-114-8587	Sequence 8587, Ap
33	513	33.7	1727	18	US-10-425-114-7603	Sequence 7603, Ap
34	513	33.7	1727	18	US-10-425-114-7624	Sequence 7624, Ap
35	513	33.7	1729	18	US-10-425-114-8164	Sequence 8164, Ap
36	513	33.7	1730	18	US-10-425-114-8133	Sequence 8133, Ap
37	513	33.7	1834	18	US-10-424-599-28863	Sequence 28863, A
38	511.4	33.6	1458	21	US-10-409-993-14	Sequence 14, Appl
39	511.4	33.6	1729	18	US-10-425-114-8584	Sequence 8584, Ap
40	505	33.1	1746	21	US-09-805-694B-15	Sequence 15, Appl
41	505	33.1	1746	21	US-10-922-732-15	Sequence 15, Appl
42	504	33.1	1458	9	US-09-758-652-12	Sequence 12, Appl
43	504	33.1	1458	18	US-10-684-651-12	Sequence 12, Appl
44	504	33.1	1458	19	US-10-757-074-12	Sequence 12, Appl
45	504	33.1	1458	19	US-10-757-155-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-731-375A-3
; Sequence 3, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731,375A
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-731-375A-3

Query Match	100.0%;	Score 1524;	DB 10;	Length 1524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1524;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CGGAGCAACCGAGGAGGAGAA	CGGTCGCGCTCAATCGCAGAGACTGAC	60
Db	1	CGGAGCAACCGAGGAGGAGAA	CGGTCGCGCTCAATCGCAGAGACTGAC	60
Qy	61	AATCGATTGATTCAGAGCGGCTTACATTGAGACTTGAACCCCAACACGAGGATTC	120	

Db 61 AATCGCATTTGAATCAGAGGGCGGTACTATTGAGACTTTGGAAACCCCAACCAACAGGAGTTTC 120
QY 121 GAATGGCGGGCTCGCCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTTCGTAGGCTT 180
Db 121 GAATGGCGGGCTCGCCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTTCGTAGGCTT 180
QY 181 TTCTACTCCAATGCTCCCCAGGAGATCTTTCATCAGCAAGGAAGGGGATACATTTGGGTTG 240
Db 181 TTCTACTCCAATGCTCCCCAGGAGATCTTTCATCAGCAAGGAAGGGGATACATTTGGGTTG 240
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCACACAAAGTTCGTGATCTCAG 300
QY 301 TCCCAAGACCAACAAGAGCTTCCAAGGAGAGACCAAGGCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACAAGAGCTTCCAAGGAGAGACCAAGGCCAACAGCAACGAGATAGT 360
QY 361 CACCAGAAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCCAACCGGTGTGCT 420
Db 361 CACCAGAAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCCAACCGGTGTGCT 420
QY 421 TTCTGGCTCTACAAACGACACGACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 421 TTCTGGCTCTACAAACGACACGACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
QY 481 AACAAAGACCAACGAGTTGATGATCTCCCGAGGAGTTCAATTTGGCTGGGAACCGGAG 540
Db 481 AACAAAGACCAACGAGTTGATGATCTCCCGAGGAGTTCAATTTGGCTGGGAACCGGAG 540
QY 541 CAAGAGTTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTTACCATAT 600
Db 541 CAAGAGTTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTTACCATAT 600
QY 601 AGCCCTACAGCCCGGCAAGTACGCTAGCAGCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 601 AGCCCTACAGCCCGGCAAGTACGCTAGCAGCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
QY 661 CAGCAGACCGCAGAGAACGAGCAGACGAGCAAGAGAGAAACCAAGTGGAAACATCTTC 720
Db 661 CAGCAGACCGCAGAGAACGAGCAGACGAGCAAGAGAGAAACCAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTACGCGCGGAGTTCTTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 AGCGGCTTACGCGCGGAGTTCTTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
QY 781 CAAAACTTAAGCGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGA 840
Db 781 CAAAACTTAAGCGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGA 840
QY 841 GGCTCAGAACTTTAGCCCGCAGATAGAAAGAGACGTCGCCAGCAAGAGAGGAATACGAT 900
Db 841 GGCTCAGAACTTTAGCCCGCAGATAGAAAGAGACGTCGCCAGCAAGAGAGGAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
QY 961 GGGAAATGGTATGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTGTGAGAAC 1020
Db 961 GGGAAATGGTATGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTGTGAGAAC 1020
QY 1021 AGATCCCTCAGATCTACAAACCTCAAGCTGGTTTCACTCAAAACTGCAACGATCTCAAC 1080
Db 1021 AGATCCCTCAGATCTACAAACCTCAAGCTGGTTTCACTCAAAACTGCAACGATCTCAAC 1080
QY 1081 CTTCTAATACTTGGTGGCTTGGACCTTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACTTGGTGGCTTGGACCTTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTGTCGCTCACTACAAACCAAGCAGACAGCATCATATATGATTTGAGGGGACGG 1200
Db 1141 TTGTTGTCGCTCACTACAAACCAAGCAGACAGCATCATATATGATTTGAGGGGACGG 1200

RESULT 2

US-10-228-806-5
; Sequence 5, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-228-806-5

Query Match 100.0%; Score 1524; DB 14; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCAGCAACCGGAGGAGAACCGTGCACAGTTCACAGCGCTCAATGCGCAGAGACCTGAC 60
Db 1 CGGCAGCAACCGGAGGAGAACCGTGCACAGTTCACAGCGCTCAATGCGCAGAGACCTGAC 60
QY 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTC 120
Db 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTC 120
QY 121 GAATGGCGGGCTCGCCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTTCGTAGGCTT 180
Db 121 GAATGGCGGGCTCGCCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTTCGTAGGCTT 180
QY 181 TTCTACTCTCAATGCTCTCCCGAGGAGATCTTTCATCCAGCAAGGAAGGGGATACATTTGGGTTG 240
Db 181 TTCTACTCTCAATGCTCTCCCGAGGAGATCTTTCATCCAGCAAGGAAGGGGATACATTTGGGTTG 240
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCACACAAAGTTCGTGATCTCAG 300
QY 301 TCCCAAGACCAACCAAGAGCTTCCAAGGAGAGACCAAGGCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACCAAGAGCTTCCAAGGAGAGACCAAGGCCAACAGCAACGAGATAGT 360


```
Qy 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGCATCTCATTTGCAGTTCCCGCGGTGTGCT 420
Db 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGCATCTCATTTGCAGTTCCCGCGGTGTGCT 420
Qy 421 TTCTGGCTCTCAACAGCACGACGACACTGATGTTGTTGCTTCTCTTACTTGACACCAAC 480
Db 421 TTCTGGCTCTCAACAGCACGACGACACTGATGTTGTTGCTTCTCTTACTTGACACCAAC 480
Qy 481 AACAAACGAAACACAGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG 540
Db 481 AACAAACGAAACACAGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG 540
Qy 541 CAAGAGTCTTAAAGGTACCGACCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAAGGTACCGACCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA 660
Db 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA 660
Qy 661 CAGCACAGCCGACAGAAACGAGCAGGACAAAGAAAGAAACAAAGAGTGGAAACATCTTC 720
Db 661 CAGCACAGCCGACAGAAACGAGCAGGACAAAGAAAGAAACAAAGAGTGGAAACATCTTC 720
Qy 721 AGCGGCTTCAGCCCGGAGTTCTTGGAAACAGCCTTCCAGTTTGACGACAGACATAGTG 780
Db 721 AGCGGCTTCAGCCCGGAGTTCTTGGAAACAGCCTTCCAGTTTGACGACAGACATAGTG 780
Qy 781 CAAAACTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACAGTCAGGGGA 840
Db 781 CAAAACTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACAGTCAGGGGA 840
Qy 841 GGCCTCAGAACTTTGAGCCAGATAGAAAGAGACGTGCCGACGAAGAGGAAATACGAT 900
Db 841 GGCCTCAGAACTTTGAGCCAGATAGAAAGAGACGTGCCGACGAAGAGGAAATACGAT 900
Qy 901 GAAGATGAATATGAATAGCATGAAGAGATAGAGGGGTGCGAGGGGAAGCAGAGGAGG 960
Db 901 GAAGATGAATATGAATAGCATGAAGAGATAGAGGGGTGCGAGGGGAAGCAGAGGAGG 960
Qy 961 GGGAAATGGTATTGAAGAGACGATCTGCACCCAGTGTCTAAAGAAACATTTGGTAGAAC 1020
Db 961 GGGAAATGGTATTGAAGAGACGATCTGCACCCAGTGTCTAAAGAAACATTTGGTAGAAC 1020
Qy 1021 AGATCCCTGACATCTCAACCCCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC 1080
Db 1021 AGATCCCTGACATCTCAACCCCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC 1080
Qy 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGAATATGGAAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGAATATGGAAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTTGTGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGG 1200
Db 1141 TTGTTTGTGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGG 1200
Qy 1201 GCTCAGCTGCAAGTCGTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGCTGCAAGTCGTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTCAAGTGTGTTGGTGGCCACAGAACTTCGCGCTCGCTGGAAAGTCCGAGGCGGAAC 1320
Db 1261 GGTCAAGTGTGTTGGTGGCCACAGAACTTCGCGCTCGCTGGAAAGTCCGAGGCGGAAC 1320
Qy 1321 TTCCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGACAGATAGCAACCTCGCGGTGAA 1380
Db 1321 TTCCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGACAGATAGCAACCTCGCGGTGAA 1380
Qy 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAATTCATATGGCCTCCAAAGG 1440
Db 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAATTCATATGGCCTCCAAAGG 1440
Qy 1441 GAGCAGGCAAGGAGCTTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
```

```
Db 1441 GAGCAGGCAAGGAGCTTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Qy 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524

RESULT 3
US-10-100-303A-89
; Sequence 89, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-89

Query Match 100.0%; Score 1524; DB 17; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAGCAACCGGAGGAGAACGCGTGCAGTTCCAGCGCTCAATGCGAGAGACCTTGAC 60
Db 1 CGGAGCAACCGGAGGAGAACGCGTGCAGTTCCAGCGCTCAATGCGAGAGACCTTGAC 60
Qy 61 AATCGCATTTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGGATTC 120
Db 61 AATCGCATTTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGGATTC 120
Qy 121 GAATCGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACCGCTTCGTAGGCT 180
Db 121 GAATCGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACCGCTTCGTAGGCT 180
Qy 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCAGCAAGAAAGGGGATATCTTTGGGTG 240
Db 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCAGCAAGAAAGGGGATATCTTTGGGTG 240
Qy 241 ATATTCCTCGTGTGCTAGACACTATGAAAGCGCTTACACAAAGTCTCGATCTCAG 300
Db 241 ATATTCCTCGTGTGCTAGACACTATGAAAGCGCTTACACAAAGTCTCGATCTCAG 300
Qy 301 TCCCAAGACACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Qy 361 CACCAAGAGTGCACCGTTTCGATGAGGGTGTCTCATTCAGAGTTCCCAACGGGTGTGCT 420
Db 361 CACCAAGAGTGCACCGTTTCGATGAGGGTGTCTCATTCAGAGTTCCCAACGGGTGTGCT 420
Qy 421 TTCTGGCTCTACACGACCAACGACACTGATGTTGTTGCTTCTTACTTGACACCAAC 480
Db 421 TTCTGGCTCTACACGACCAACGACACTGATGTTGTTGCTTCTTACTTGACACCAAC 480
Qy 481 AACAAACGAAACACAGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG 540
Db 481 AACAAACGAAACACAGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACACGGAG 540
Qy 541 CAAGAGTCTTAAAGGTACCGACCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAAGGTACCGACCAACAAAGCAGACCAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA 660
```

Db 601 AGCCCATACAGCCCGCAAAAGTACGCTTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
Qy 661 CAGCACAGCCGCGAGAGAAACGAGCAGGACAAAGAAAGAAACAAAGAGTGAAGAAATCTTC 720
Db 661 CAGCACAGCCGCGAGAGAAACGAGCAGGACAAAGAAAGAAACAAAGAGTGAAGAAATCTTC 720
Qy 721 AGCGGCTTACGCGCGGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 AGCGGCTTACGCGCGGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Qy 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTGACAGTGAAGGGA 840
Db 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTGACAGTGAAGGGA 840
Qy 841 GGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGAATACGAT 900
Db 841 GGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGAATACGAT 900
Qy 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGGCGTGGCAGGGGAAGCAGAGGCAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGGCGTGGCAGGGGAAGCAGAGGCAGG 960
Qy 961 GGGAAATGGTATTAAGAGAGAGATCTGCACCGCAAGTCTGCTAAAAGAAACATTTGGTAGAAAC 1020
Db 961 GGGAAATGGTATTAAGAGAGAGATCTGCACCGCAAGTCTGCTAAAAGAAACATTTGGTAGAAAC 1020
Qy 1021 AGATCCCTCAGACATCTACAACCTCAAGCTGGTTCACTCAAAACTGCAACAGATCTCAAC 1080
Db 1021 AGATCCCTCAGACATCTACAACCTCAAGCTGGTTCACTCAAAACTGCAACAGATCTCAAC 1080
Qy 1081 CTTCTAATACCTTAGTGGCTTGGACCTTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACCTTAGTGGCTTGGACCTTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTGTCTCCTCACTACACACCAAGCAGACAGCATCATATATCGATTGAGGGGACGG 1200
Db 1141 TTGTTGTCTCCTCACTACACACCAAGCAGACAGCATCATATATCGATTGAGGGGACGG 1200
Qy 1201 GCTCAGCTGAAGTCGTGGAGCAGCAACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGCTGAAGTCGTGGAGCAGCAACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTCAAGTGTCTGTGGTGCACAGAACTTCGCGGTGCTGGAAAGTCCGAGAGCGAGAAC 1320
Db 1261 GGTCAAGTGTCTGTGGTGCACAGAACTTCGCGGTGCTGGAAAGTCCGAGAGCGAGAAC 1320
Qy 1321 TTCGAATACGTGCATCAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCGGTGAA 1380
Db 1321 TTCGAATACGTGCATCAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCGGTGAA 1380
Qy 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGTGGTTGCAATTCATATGGCCTCCAAAGG 1440
Db 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGTGGTTGCAATTCATATGGCCTCCAAAGG 1440
Qy 1441 GAGCAGCAAGGAGCGCTTAAGAAACAACACCCCTTCAAGTTCCTCGTTCACCGCTCTCAG 1500
Db 1441 GAGCAGCAAGGAGCGCTTAAGAAACAACACCCCTTCAAGTTCCTCGTTCACCGCTCTCAG 1500
Qy 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524

RESULT 4
US-10-728-323-3
; Sequence 3, Application US/10728323
; Publication No. US20040208894A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh

; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0232
; CURRENT APPLICATION NUMBER: US/10/728,323
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-10-728-323-3

Query Match 100.0%; Score 1524; DB 20; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCAGCAACCGGAGGAGAACGCGTCCAGTTCAGCGCCTCAATGCGCAGAGACCTTGAC 60
Db 1 CGGCAGCAACCGGAGGAGAACGCGTCCAGTTCAGCGCCTCAATGCGCAGAGACCTTGAC 60
Qy 61 AATCGCATTTGAATCAGAGGCGGTTTACATTGAGACTTTGGAACCCCAACCAACAGAGATTTC 120
Db 61 AATCGCATTTGAATCAGAGGCGGTTTACATTGAGACTTTGGAACCCCAACCAACAGAGATTTC 120
Qy 121 GAATGCGCGCGGTGCGCCTCTCTCGTTAGTCTCTCGCGCGCAACGCGCCTTGTTAGGCCT 180
Db 121 GAATGCGCGCGGTGCGCCTCTCTCGTTAGTCTCTCGCGCGCAACGCGCCTTGTTAGGCCT 180
Qy 181 TTCTACTCTCAATGCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGAGTACTTTGGGTTG 240
Db 181 TTCTACTCTCAATGCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGAGTACTTTGGGTTG 240
Qy 241 ATATTCCTCGTTGTCCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCTCGTTGTCCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
Qy 301 TCCCAAGACCAACCAAGAGCGTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACCAAGAGCGTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Qy 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT 420
Db 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT 420
Qy 421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGTTGTTCTTCTTACTGACACCAAC 480
Db 421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGTTGTTCTTCTTACTGACACCAAC 480
Qy 481 AACACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCATATTTGGCTGGGAACACGGAG 540
Db 481 AACACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCATATTTGGCTGGGAACACGGAG 540
Qy 541 CAAGAGTCTTAAGGTGACGACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAGGTGACGACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAAAGTCAAGAGAGCGTGAATTTTAGCCCTCGAGGA 660
Db 601 AGCCCATACAGCCCGCAAAAGTCAAGAGAGCGTGAATTTTAGCCCTCGAGGA 660
Qy 661 CAGCACAGCCGCGAGAGAACGAGCAGGACAAAGAAAGAAACAAAGGTGGAAACATCTTC 720
Db 661 CAGCACAGCCGCGAGAGAACGAGCAGGACAAAGAAAGAAACAAAGGTGGAAACATCTTC 720
Qy 721 AGCGGCTTCAACGCGGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 AGCGGCTTCAACGCGGAGTTCCTGGNAACAAGCCTTCAGGTTGACGACAGACAGATAGTG 780
Qy 781 CAAAACCTTAAGAGGCGAGACCGGAGAGTGAAGAGGGAGGCCATTTGTGACGTGAGGGGA 840

Db 781 CAAAACCTTAGAGCCGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
Qy 841 GGCTCTCAGATCTTTGAGCCCGAGATAGAAAGAGAGCGTCCGACGAAAGAGAGGAATACGAT 900
Db 841 GGCTCTCAGATCTTTGAGCCCGAGATAGAAAGAGAGAGCGTCCGACGAAAGAGAGGAATACGAT 900
Qy 901 GAAGATGAATATGATACGATAGAGAGATAGAAAGGCGTGGCAGGGAAGCAGAGGAGG 960
Db 901 GAAGATGAATATGATACGATAGAGAGATAGAAAGGCGTGGCAGGGAAGCAGAGGAGG 960
Qy 961 GGGAAATGGTATTAAGAGAGCGATCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
Db 961 GGGAAATGGTATTAAGAGAGCGATCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
Qy 1021 AGATCCCTGACATCTTCAACCCCTCAAGCTGGTTCACTCAAACTGCCAAGATCTCAAC 1080
Db 1021 AGATCCCTGACATCTTCAACCCCTCAAGCTGGTTCACTCAAACTGCCAAGATCTCAAC 1080
Qy 1081 CTTCTAATACCTTAGTGGCTTGACCTAGTGTGCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACCTTAGTGGCTTGACCTAGTGTGCTGAATATGGAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTTGTGCTCACTACAACCAACGACGACAGCATATATCGATTGAGGGACGG 1200
Db 1141 TTGTTTGTGCTCACTACAACCAACGACGACAGCATATATCGATTGAGGGACGG 1200
Qy 1201 GCTCAGTGAAGTCTGTGAGCAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGTGAAGTCTGTGAGCAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTCAAGTGTGTGTGTCACAGAACTTCGCGTGTGTCGGAAGTCCGACGAGGAGAAC 1320
Db 1261 GGTCAAGTGTGTGTGTCACAGAACTTCGCGTGTGTCGGAAGTCCGACGAGGAGAAC 1320
Qy 1321 TTGGAATACGTGTCATTCAGACAGACTCAAGGCCAGCAGTACCAACCTCGCGGTGAA 1380
Db 1321 TTGGAATACGTGTCATTCAGACAGACTCAAGGCCAGCAGTACCAACCTCGCGGTGAA 1380
Qy 1381 AACTCCGTATAGATAACCTGCGGAGGAGTGTGTCAAATTCATATGSCCTCCAAAGG 1440
Db 1381 AACTCCGTATAGATAACCTGCGGAGGAGTGTGTCAAATTCATATGSCCTCCAAAGG 1440
Qy 1441 GAGCAGCAAGGAGCTTAAGAACCAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Db 1441 GAGCAGCAAGGAGCTTAAGAACCAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Qy 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524

RESULT 5

US-10-728-051-3
; Sequence 3, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3

; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-728-051-3
Query Match 100.0%; Score 1524; DB 20; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCACAAACCGAGGAGAAACGCGTCCAGTTCCAGCGCTCAATGCGAGAGACCTGAC 60
Db 1 CGGCACAAACCGAGGAGAAACGCGTCCAGTTCCAGCGCTCAATGCGAGAGACCTGAC 60
Qy 61 AATCGCATTAATCAGAGGGCGGTTACATTGAGACTTTGGAACCCCAACCAACGAGAGTTC 120
Db 61 AATCGCATTAATCAGAGGGCGGTTACATTGAGACTTTGGAACCCCAACCAACGAGAGTTC 120
Qy 121 GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 121 GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Qy 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCAATCCAGAAAGGAGGATCTTTGGGTG 240
Db 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCAATCCAGAAAGGAGGATCTTTGGGTG 240
Qy 241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCCTCACACAAGGTCGTTCGATCTCAG 300
Db 241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCCTCACACAAGGTCGTTCGATCTCAG 300
Qy 301 TCCCAAGAACCCACCAAGAGCTCTCCAAAGAGAGACCAAGCCCAACGAGATAGT 360
Db 301 TCCCAAGAACCCACCAAGAGCTCTCCAAAGAGAGACCAAGCCCAACGAGATAGT 360
Qy 361 CACCAGAGGTGACCGTTTCCGATGAGGCTGATCTCATTTGCGATTTCCCGGTGTGCT 420
Db 361 CACCAGAGGTGACCGTTTCCGATGAGGCTGATCTCATTTGCGATTTCCCGGTGTGCT 420
Qy 421 TTCTGCTCTCAACGACCACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
Db 421 TTCTGCTCTCAACGACCACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
Qy 481 AACACGACCAACCGTTCGATCAGTTCCCGAGGAGATTCAAATTTGCTGGACACGGAG 540
Db 481 AACACGACCAACCGTTCGATCAGTTCCCGAGGAGATTCAAATTTGCTGGAAACGAGG 540
Qy 541 CAAGAGTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAGTCAAGTACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 601 AGCCCATACAGCCCGCAAGTCAAGTACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Qy 661 CAGCAGCCCGCAGAGAACGAGCAGCAAGAGAGAAACGAAAGTGGAAACATCTTC 720
Db 661 CAGCAGCCCGCAGAGAACGAGCAGCAAGAGAGAAACGAAAGTGGAAACATCTTC 720
Qy 721 AGCGGCTTCAAGCGGAGTCTTGGAAACAGCTTCCAGGTTGACACAGAGATAGTG 780
Db 721 AGCGGCTTCAAGCGGAGTCTTGGAAACAGCTTCCAGGTTGACACAGAGATAGTG 780
Qy 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGGAGGACCATTTGTGACAGTGGGA 840
Db 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGGAGGACCATTTGTGACAGTGGGA 840
Qy 841 GGCCTCAGATCTTGAAGCCAGATAGAAAGAGACCTGCGCAGCAAGAGAGGAATACGAT 900
Db 841 GGCCTCAGATCTTGAAGCCAGATAGAAAGAGACCTGCGCAGCAAGAGAGGAATACGAT 900
Qy 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGCTGGCAGGGAAGCAGGACG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGCTGGCAGGGAAGCAGGACG 960

```
QY 961 GGAATGGTATTGAAGAGACGATCTGCAACGCAAGTCTAAAAAGAACATTTGGTAGAAAC 1020
Db |||||
961 GGAATGGTATTGAAGAGACGATCTGCAACGCAAGTCTAAAAAGAACATTTGGTAGAAAC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCTAATAAATCGCAACGATCTCAAC 1080
Db |||||
1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCTAATAAATCGCAACGATCTCAAC 1080
QY 1081 CTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA 1140
Db |||||
1081 CTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTTGTGGCTCACTACAAACCGCACACAGCATCATATATCGATTGAGGGACGG 1200
Db |||||
1141 TTGTTTGTGGCTCACTACAAACCGCACACAGCATCATATATCGATTGAGGGACGG 1200
QY 1201 GCTCAGTGAAGTCTGGAGACGAAACGGCAACAGAGTGTAACAAGAGCTTCAAGAG 1260
Db |||||
1201 GCTCAGTGAAGTCTGGAGACGAAACGGCAACAGAGTGTAACAAGAGCTTCAAGAG 1260
QY 1261 GGTCAAGTCTTGTGGTCCACAGAACTTCGCCGTGCTGGAAAGTCCAGAGCGAGAAC 1320
Db |||||
1261 GGTCAAGTCTTGTGGTCCACAGAACTTCGCCGTGCTGGAAAGTCCAGAGCGAGAAC 1320
QY 1321 TTGGAATAGTGGCAATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCCGTGAA 1380
Db |||||
1321 TTGGAATAGTGGCAATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCCGTGAA 1380
QY 1381 AACTCGTCTAGATTAACCTGCCGAGAGAGTGTTGCAAAATTCATATATGGCTTCAAAGG 1440
Db |||||
1381 AACTCGTCTAGATTAACCTGCCGAGAGAGTGTTGCAAAATTCATATATGGCTTCAAAGG 1440
QY 1441 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCCACCGTCTCAG 1500
Db |||||
1441 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
Db |||||
1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
```

RESULT 6

```
US-10-899-551-5
; Sequence 5, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burke, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899, 551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: species Arachis hypogaea
US-10-899-551-5
```

```
Query Match 100.0%; Score 1524; DB 21; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGGCAGCAACCGGAGGAGAACGGTGCCAGTTCACAGCGCTCAATCGCGAGAGACCTGAC 60
Db |||||
1 CGGCAGCAACCGGAGGAGAACGGTGCCAGTTCACAGCGCTCAATCGCGAGAGACCTGAC 60
QY 61 AATCGCATTTGAATCAGAGGCGGTTACATTTGAGACTTTGGAACCCCAACACAGAGATTTC 120
Db |||||
```

```
Db 61 AATCGCATTTGAATCAGAGGCGGTTACATTTGAGACTTTGGAACCCCAACACAGAGATTTC 120
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCGCCCTCTGTTAGSCCT 180
Db |||||
121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCGCCCTCTGTTAGSCCT 180
QY 181 TTCTACTCTCAATGCTCTCCAGAGAGATCTTTCAATCCAGCAAGAGAGGGGATACATTTTGGGTTG 240
Db |||||
181 TTCTACTCTCAATGCTCTCCAGAGAGATCTTTCAATCCAGCAAGAGAGGGGATACATTTTGGGTTG 240
QY 241 ATATTCCCTGGTTGCTCTAGACATCATATGAAGAGCTTCAACAAGGTCGTCTGATCTCAG 300
Db |||||
241 ATATTCCCTGGTTGCTCTAGACATCATATGAAGAGCTTCAACAAGGTCGTCTGATCTCAG 300
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db |||||
301 TCCCAAGACCAACCAAGAGCTCTCCAAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
QY 361 CACCAAGAGGTGCAACCGTTTCGATGAGGAGTCTCTATTGCAAGTTCCTTACTGACACCAAC 420
Db |||||
361 CACCAAGAGGTGCAACCGTTTCGATGAGGAGTCTCTATTGCAAGTTCCTTACTGACACCAAC 420
QY 421 TTCTGGCTCTCAACGACACGACATCTGATGTTGTTGCTGTTCTTCTACTGACACCAAC 480
Db |||||
421 TTCTGGCTCTCAACGACACGACATCTGATGTTGTTGCTGTTCTTCTACTGACACCAAC 480
QY 481 AACACGACCAACCGCTTGTATCCCGAGAGATTCATTTGCTGGGAAACGCGAG 540
Db |||||
481 AACACGACCAACCGCTTGTATCCCGAGAGATTCATTTGCTGGGAAACGCGAG 540
QY 541 CAAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db |||||
541 CAAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
QY 601 AGCCCATACAGCCCGCAAGTCTAGACAAAGAGAGCTGAAATTTAGCCCTCGAGGA 660
Db |||||
601 AGCCCATACAGCCCGCAAGTCTAGACAAAGAGAGCTGAAATTTAGCCCTCGAGGA 660
QY 661 CAGCAGCGCGCAGAGAACGAGCAGACAAAGAGAGAAACGAAAGTGGAAACATCTTC 720
Db |||||
661 CAGCAGCGCGCAGAGAACGAGCAGACAAAGAGAGAAACGAAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTTCAGCCCGGAGTCTCTGGAAACAAGCTTTCAGGTTGACGACAGACAGATAGTG 780
Db |||||
721 AGCGGCTTTCAGCCCGGAGTCTCTGGAAACAAGCTTTCAGGTTGACGACAGATAGTG 780
QY 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGGAGCCATTTGTGACGTGAGGGA 840
Db |||||
781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGGAGCCATTTGTGACGTGAGGGA 840
QY 841 GGCCTCAGAACTCTTGAGCCCGAGATAGAAAGACGCTGCCGACGAAAGAGAGAAATACGAT 900
Db |||||
841 GGCCTCAGAACTCTTGAGCCCGAGATAGAAAGACGCTGCCGACGAAAGAGAGAAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGCGCAGGGGAGCAGAGCGAG 960
Db |||||
901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGCGCAGGGGAGCAGAGCGAG 960
QY 961 GGGAAATGTTTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
Db |||||
961 GGGAAATGTTTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAAC 1080
Db |||||
1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAAC 1080
QY 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTTACAGGAATGCA 1140
Db |||||
1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTTACAGGAATGCA 1140
QY 1141 TTGTTTGTGGCTCACTACAAACCGCACACAGCATCATATATCGATTGAGGGACGG 1200
Db |||||
1141 TTGTTTGTGGCTCACTACAAACCGCACACAGCATCATATATCGATTGAGGGACGG 1200
```

Qy	1201	GCTCAGTGC	AAGTCGTGG	CAGCAACGG	CAACAGAGTGT	TACGACGAGG	AGCTTCAAG	AG	1360
Db	1201	GCTCAGTGC	AAAGTCGTGG	CAGCAACGG	CAACAGAGTGT	TACGACGAGG	AGCTTCAAG	AG	1360
Qy	1261	GGTCACGTGC	TTGTGTGGTGC	CACAGAACTT	TCGCGTGC	GTGAAAGTCC	CAGAGCGGAAC		1320
Db	1261	GGTCACGTGC	TTGTGTGGTGC	CACAGAACTT	TCGCGTGC	GTGAAAGTCC	CAGAGCGGAAC		1320
Qy	1321	TTTCAATACT	CGTGGCATTCA	GACAGACAG	ACTCAAGGCCC	AGCATAGACCA	ACCTTCGCGGTGA		1380
Db	1321	TTTCAATACT	CGTGGCATTCA	GACAGACAG	ACTCAAGGCCC	AGCATAGACCA	ACCTTCGCGGTGA		1380
Qy	1381	AACTCCGT	CATAGATAA	CTCCCGAGG	AGGTGGTTG	CAAAATTCAT	TATGGCTCCAA	AGG	1440
Db	1381	AACTCCGT	CATAGATAA	CTCCCGAGG	AGGTGGTTG	CAAAATTCAT	TATGGCTCCAA	AGG	1440
Qy	1441	GAGCAGGCA	AGGCAAGT	TAAAGAA	CAACCCCTT	CAAGTTCTT	CGTTCCACCGT	CTCAG	1500
Db	1441	GAGCAGGCA	AGGCAAGT	TAAAGAA	CAACCCCTT	CAAGTTCTT	CGTTCCACCGT	CTCAG	1500
Qy	1501	CAGTCTCG	GAGGGCTGT	GGCTTAA					1524
Db	1501	CAGTCTCG	GAGGGCTGT	GGCTTAA					1524

RESULT 7

```

US-10-358-324-4
; Sequence 4, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: BENTZEN, CHARLES J.
; APPLICANT: KONAN, KOFEI N'DA
; APPLICANT: VQUEEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-958-324-4

```

Query Match	91.3%	Score 1391.2	DB 21	Length 1853
Best Local Similarity	95.4%	Pred. No. 0		
Matches 1454	Conservative 0	Mismatches 68	Indels 2	Gaps 2
Qy	2	GGCAGCAACCGGAGGAGAAACGCGTGCAGATTTCACAGCGCTCAATGCGCAGAGACTGACA	61	
Db	71	GGCAGACCGGAGGAGGAGATGCGTGCAGTTCAGCGGCTCAATGCGCAGAGACTGACA	130	
Qy	62	ATCGCATTTGAATCAGAGGGCGGTTACATTTGAGATTGGAAACCCCAACAACACGAGGTTTCG	121	
Db	131	ACCCGATTTGAATCGGAGGGCGGTTTACATTTGAGATTGGAAACCCCAACAACACGAGGTTTCG	190	
Qy	122	AATCGCGCGGCGTGCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTCGTAGGCGCTT	181	
Db	191	AATCGCGCGGCGTGCCTCTCTCGCTTAGTTCCTCGCGCGCAACGCCCTTCGTAGGCGCTT	250	
Qy	182	TCTACTCCAATGCTCCCGCAGAGATCTTCATCAGCAAGGAAGGGGATCTTTGGGTTGA	241	
Db	251	TCTACTCCAATGCTCCCGCAGAGATCTTCATTCAGCAAGGAAGGGGATCTTTGGGTTGA	310	
Qy	242	TATTCCCTGGTGTTCCTAGACATATGAAGAGCGCTCACAACAAGGTCGTGCATCTCAGT	301	

Db 1390 TTCGAATACGTGGCAATTCGAAGACAGATTCAAGGCCAGCATAGCCAACTTTGCGGTGAA 1449
QY 1381 AACTCCGTATAGATAAATCTGCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 1450 AACTCTTTCATAGATAAATCTGCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAG 1509
QY 1441 GAGCAGCGAAGGAGCTTAAAGAAACAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Db 1510 GAGCAGCGAAGGAGCTTAAAGAAACAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1569
QY 1501 CAGTCTCCGAGGCTGTGCTTAA 1524
Db 1570 CAGTCTCCGAGGCTGTGCTTAA 1593

RESULT 8

US-09-758-652-13
; Sequence 13, Application US/09758652
; Patent No. US20010011377A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSER: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898

COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-758-652-13

Query Match 34.0%; Score 517.4; DB 9; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.7e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;
QY 5 AGCAACCGGAGGAACGGGTCCAGTTCAGCGCTCAATGCGCAGAGACCTTGACAATC 64
Db 71 AGCAGCCACAGCAAAACGAGTGCAGATCCAAAGCCTCAATGCCCTTAAACCGGATAAC 130
QY 65 GCATTGAATCAGAGGGGGTTACATTTGAGACTTGGAAACCCACACACCGAGGTTCCGAAT 124
Db 131 GTATAGAGTCAGAAGGTGGCTTCATTTGAGACATGGAAACCCCTTACAAACAAGCCATTCAGT 190

QY 125 GGGCGGGGTGCGCCCTCTCTCGCTTAGTCTCCGCCGCAACGCCCTTCTGTAAGCCTTTCT 184
Db 191 GTGCCGGTGTGGCCCTCTCTCGCTCACCCCTCAACGGCAACGCCCTTCCGACAGCCTTCT 250
QY 185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACCTTTGGTGTGATAT 244
Db 251 ACACCAACGCTCCCGAGGAGATCTACATCCAAACAAGGTAGTGGTATTTTGGCATGATAT 310
QY 245 TCCCTGGTGTCTAGACACTATGAAGAGCCTCACACACAAGGTGCTCGATCTCAGTCCC 304
Db 311 TCCCGGGTGTCTTAGCACTTTGAAGAGCCT----- 342
QY 305 AAAGACCACCAAGAGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGTCACC 364
Db 343 -----CAACAAAAGGACAAAGCAGCAGGCCCCCAAGACCGTCACC 382
QY 365 AGAAGGTGACCGGTTTCGATGAGGGTGATCTCATTTGCGAGTTCCACCGGTGTGCTTTCT 424
Db 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTCGAGTGCACACCGGTTTGCATACT 442
QY 425 GGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACAACA 484
Db 443 GGATGTACAAACATGAAGACACTCCTGTTGTTGCGGTTTCTCTTATTTGACACCAACAGCT 502
QY 485 ACGACAACCCAGCTTGATCAGTTCCTCCAGGAGATTCAATTTTGGCTGGGAACACGGAGCAAG 544
Db 503 TCCAGAACCGCTCGACAGATGCTCTAGGAGATTCTATCTCTGCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAGAAAGAGCTTACCATATAGCC 604
Db 563 AGTTCTTACAGTATCAGCCACAGAAAGCAGCAAG----- 595
QY 605 CATACAGCCGCAAGTCTAGCCTTAGCAAGAAGAGCGTGAAATTTAGCCCTCGAGGACAGC 664
Db 596 -----GAGGTACTC 604
QY 665 ACAGCCGAGAGACGAGCAGGACGAAGAAGAAAGAAAGAGTGGAAACATCTTTCAGCG 724
Db 605 AAAGCCAGAAAGGAAGGCTCAGCAAGAAGAAAGAAAGAAAGAGGAGCAGCATATTGAGTG 664
QY 725 GCTTCAGCGCGAGTTCCTGGAAACAGCTTCCAGGTTGACGACAGACAGATAGTCAAA 784
Db 665 GCTTCGCGCGGAAATTTTGGAAACATGCGTTTC---GTCGTGGACAGGAGATAGTGAGAA 721
QY 785 ACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGAGGAGCCATTTCTGACAGTGAAGGAGGCC 844
Db 722 AGCTCAAGGTGAGAACGAGAGGAGAGAGAGGTTGCCATTTGTGACAGTGAAGAGGAGTTC 781
QY 845 TCAGAAATCTTGAGCCCGAGATAGAAAGAGACGTCGCCGACGAAGAAGAGGAAATACGATGAAG 904
Db 782 TCAGCGTGATAAGCCACCCACGGAAGAGCAGCAACAAGACCCGAGGAAGAGGAGAAAGC 841
QY 905 ATGAATATGAATACGATGAAGAGGATAGNAGCGTGGCGAGGAGAGCAGAGCAGGGGGA 964
Db 842 CAGATTGTGACAGAAAGACAAACAT-----TGCCAAAGCCAAAGACAGAA 886
QY 965 ATGGTATTGAAGAGACGATCTGACCGCAAGTGCTAAAGAAAGAAACATTTGTAGAAACAGAT 1024
Db 887 ATGGCATTTGACAGACCATTTGCAACAATGAGACTTTGCCACACCAATTTGGCCAGACTTCA 946
QY 1025 CCCCTGACATCTCAACCCCTCAAGTGGTTCTACTCAAAATCTGCCAACGATCTCAACCTTTC 1084
Db 947 CACCTTGACATCTTCAACCCCTCAAGTGGTAGCATCACAAACCGCTACCAGCCTCGACTTCC 1006
QY 1085 TAACTACTAGTGGCTTGGACCTAGTGTGATATAGGAAATCTCTACAGGATGCAATGTT 1144
Db 1007 CAGCCCTCTCGTGGGTCAAACTCAGTGGCCAGTTTGGATCTACTCCGCAAGATGCTATGT 1066
QY 1145 TTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTTGAGGGGACGGGCTC 1204
Db 1067 TCGTGCCACACTACAAACCTGAAACGCAACAGCATATATACCATTTGATGACGGGCAT 1126

QY 1205 ACGTGCAAGTCTGTGACAGCAACCGCAAGGTGTACGAGGAGCTTCAAGAGGGTC 1264
 DB 1127 TGGTACAAAGTGTGAATGCAATGGTGAGAGAGTGTGATGGAGAGCTGCAAGAGGAC 1186
 QY 1265 ACGTGCTGTGGTCCCAAGAACTTCCCGCTGCTGGAAAGTCCAGAGGAGAACTTCG 1324
 DB 1187 AGGTGTTAAATGTGCCCAAACTTTCGGGTGGCTGCAAGATCAAGAGGAGCAACTTCG 1246
 QY 1325 AATACGTGGATTCAAGCAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAACT 1384
 DB 1247 AGTATGTTCAATCAAGACCAATGATAGACCCCTCGATCGGCAACTTTCAGAGGTGCAAACT 1306
 QY 1385 CCGTCATAGATAAACCCTCCGAGGAGGTGGTGGCAAAATTCATATGGCTCCAAAGGGAGC 1444
 DB 1307 CATTGTTGAACGATTGCGGAGGAGTGTTCAGCAAACTTTTAACTAAGGAGGAGC 1366
 QY 1445 AGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTCTTGGTTCCACCGTCTCAGCAGT 1504
 DB 1367 AGGCCAGCAGTCAAGAAACAAACCCCTTTCAGCTTCTGGTTCCACCTAAGGAGTCTC 1426
 QY 1505 CTCCGAGGCTGTGGCTTA 1523
 DB 1427 AGAGGAGGTGTGGCTTA 1445

RESULT 9

US-10-684-651-13
 ; Sequence 13, Application US/10684651
 ; Publication No. US20040064858A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; GARY MICHAEL FADER
 ; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 ; CLASSES OF SOYBEAN SEED
 ; PROTEIN GENES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 ; STREET: 1007 MARKET STREET
 ; CITY: WILMINGTON
 ; STATE: DELAWARE
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 19898
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95
 ; SOFTWARE: MICROSOFT WORD 7.0A
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/684,651
 ; FILING DATE: 14-Oct-2003
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/758,652
 ; FILING DATE: 11-Jan-2001
 ; APPLICATION NUMBER: 60/019,940
 ; FILING DATE: JUNE 14, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: LYNNE M. CHRISTENBURY
 ; REGISTRATION NUMBER: 30,971
 ; REFERENCE/DOCKET NUMBER: BB-1071-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 302-992-5481
 ; TELEFAX: 302-773-0164
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1446 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
 US-10-684-651-13

Query Match 34.0%; Score 517.4; DB 18; Length 1446;
 Best Local Similarity 62.8%; Pred. No. 1.7e-154; Indels 144; Gaps 4;
 Matches 954; Conservative 0; Mismatches 421;
 QY 5 AGCAACCGGAGGAGAAACGGCTGCCAGTTCAGGCGCTCAATGGCGAGAGACCTTGACAATC 64
 DB 71 AGCAGCCACAGCAAAACGAGTGCAGATCCAAACGCTCAATGCCCTAAACCGGATAACC 130
 QY 65 GCATTGAATCAGAGGGGGTTACATTGAGACTTTGGAAACCCCAACACAGAGAGTTGCAAT 124
 DB 131 GTATAGAGTCAGAGGTGGCTTCAATTGAGACATGGAACCCCTAAACAACAGCCATTCAGT 190
 QY 125 GCGCCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGGCTTTCT 184
 DB 191 GTGCCGTGTTCCTCTCTCGCTTGACCTCAACCGCAACGCGCTTCGAGACCTTCCT 250
 QY 185 ACTTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGGTGATAT 244
 DB 251 ACACCAACGCTCCAGGAGATCTACATCCAAAGGTAGTGTATTTTGGCATGATAT 310
 QY 245 TCCCTGGTTGCTTAGACACTATGAAGAGCTTCACACAAGGTGCTCGATCTCAGTCCC 304
 DB 311 TCCCGGTTGTCTTAGCACATTTGAAGAGCTT----- 342
 QY 305 AAAGACCAACAGACGCTCTCRAAGGAGAAACCAAGCCAAACAGCAACGAGATAGTCACC 364
 DB 343 -----CAACAAAAGGACAAAGCAGCAGCGCCCAAGACCGTCACC 382
 QY 365 AGAAGGTGACCGTTTCGATGAGGTGTCTCATTCAGTTCCTCCACCGGTGTGCTTTCT 424
 DB 383 AGAAGATCTATCACTTCAGAGAGGGTGTATTGATTGAGTGCAGTGCACACCGGTTTGCATACT 442
 QY 425 GGCTTACAAACGACACGACACTGTGTTGCTGCTTCTTCTTACGACACCAACAACA 484
 DB 443 GGATGTACAAATGAAGACACTCTCTGTTGTTGCTGCTTCTTATTTGACACCAACAGCT 502
 QY 485 ACACAAACGAGTGTATGATTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAG 544
 DB 503 TCCAGAACCCAGCTCGACAGATGCTTAGGAGATTTCTATCTTGTGGGAACCAAGAGCAAG 562
 QY 545 AGTTCCTTAAGGTACCAGCAACAAAGCAGACAAAAGCAGACGAAGAGCTTACCATATAGCC 604
 DB 563 AGTTTCTACAGTATCAGCCACAGACGACGACG- 595
 QY 605 CATACAGCCGCAAAAGTTCAGCTTAGACAAGAGCGTGAATTTAGCCCTTCGAGGACAGC 664
 DB 596 -----GAGGTACTC 604
 QY 665 ACAGCCGACAGAAACGAGCAGGACAAAGAAACGAAACGAGGTGGAACATCTTCAGCG 724
 DB 605 AAAGCCAGAAAGGAAAGCGCTCAGCAAGAAAGAAACGAAAGGAGCGCATATTGAGTG 664
 QY 725 GCTTTCAGCGCGGAGTTCTCTGGAAACAGCCCTTCCAGGTTCAGCAGACAGATAGTGCAA 784
 DB 665 GCTTCGCGCGGAAATCTTGGAAACATGCGTTC---GTGCTGGACAGCAGATAGTGAA 721
 QY 785 ACCTAAGAGCGCAGACCGGAGGTGAAGAGAGGAGGACCAATTGTGACAGTGGGGAGGCC 844
 DB 722 AGCTACAAGGTGAGAAACGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
 QY 845 TCAGAAATCTTGAGCCCGCAGATAGAAAGAGAGCTGCCAGCAGAGAGGAGGAGGAGGAG 904
 DB 782 TCAGCGTGTATGAGCCCAACCCCAACGGAAGAGCAGCAACAAAGACCCGAGGAGAGGAGGAG 841
 QY 905 ATGAATATGAATACGATGAAGAGGTAGAGGCGTGCAGGGAAGAGCAGAGGAGGAGGAG 964
 DB 842 CAGATTGTGACGAGAAAGACAAACAT-----TCCAAAGCCAAAGCAGAA 886
 QY 965 ATGGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAACAGAT 1024
 DB 887 ATGGCATTGACGAGACCAATTTGCACATGAGACTTGCACACACATTTGGCCAGACTTCAT 946

QY 1025 CCCTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACTGCCAAGATCTCAACCTTC 1084
Db 947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCGCTACCGCTCGACTTC 1006
QY 1085 TAATACTTAGTGGCTTGGACCTAGTGTGCTGAATATGAAATCTCTACAGGAATGCATTGT 1144
Db 1007 CAGCCCTCTCGTGGCTCAAACTCAGTGCACGAGTTTGATCACTCCGCAAGAATGCTATGT 1066
QY 1145 TTGTCGCTCACTACAACCAACCGCACACAGCATCATATATCATATGATTTGAGGGGACGGGCTC 1204
Db 1067 TCGTGCCACACTACAACCTGAACGCAACAGCATATATACGCAATTTGAATGGACGGCAT 1126
QY 1205 ACCTGCAAGTCTGGACAGCAACGCGCAACAGAGTGTACGACGAGAGCTTCAAGAGGTC 1264
Db 1127 TGGTACAAGTGGTGAATTTGCAATGCTGAGAGAGTGTGATGAGAGCTGCAAGAGGGAC 1186
QY 1265 ACCTGCTTGTGGTGGCCACAGAACTTCCCGTCTGCTGGAAGTCCCGAGAGGAGAACTTCG 1324
Db 1187 AGGTGTTAATTGTGGCCACAAACTTTTGGGTGGCTGCAAGATCACAGAGCGCAACTTCG 1246
QY 1325 AATACGTGGCAATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAAACT 1384
Db 1247 AGTATGTTTCATTCAAGACCAATGATAGACCTCGATCGCAACCTTGAGGTGCAAACT 1306
QY 1385 CCCTCATAGATAACCTGCCGAGAGGTGGTTGCAAAATCATATGCGCTCCAAAGGGAGC 1444
Db 1307 CATTGTTGAACGCAATTCGCGAGGAAGTGAATTCAGCAAACTTTTAAACCTTAAGGAGGCAG 1366
QY 1445 AGCAAGGAGCTTAAGCAACAACCCCTTCAAGTTCCTTGGTTCACCGCTCTCAGCAGT 1504
Db 1367 AGCCAGGAGGAGTCAAGAAACAACCCCTTTCAGTCTCTGCTTCCCTGAGTTCACCTAAGGAGTCTC 1426
QY 1505 CTCGAGGGCTGTGGCTTA 1523
Db 1427 AGAGGAGAGTTGTGGCTTA 1445

RESULT 10
US-10-757-074-13
; Sequence 13, Application US/10757074
; Publication No. US20040139502A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10757,074
; FILING DATE: 14-Jan-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071--A

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-757-074-13

Query Match 34.0%; Score 517.4; DB 19; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.7e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

QY 5 AGCAACCGGAGAGAACGGTGCAGATTTCCAGCGCTCAATGCGGAGAGACCTGACAATC 64
Db 71 AGCAGCCACAGCAAAACGAGTGCAGATCCAAACGGCTCAATGCCCTTAAACCGGATAAC 130
QY 65 GCATTGAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTCAAT 124
Db 131 GTATAGATCAGAAGTGGCTTCAATTGAGACATGGAACCCCTAACCAAGCAATTCAGT 190
QY 125 GCGCGGGCTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCCTTCT 184
Db 191 GTGCGGGTGTGCGCTCTCTCGCTCACCCCTCAACCGCAACGCCCTTCGCAGACCTTCT 250
QY 185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACCTTTGGTGTGATAT 244
Db 251 ACACCAACGCTCCCGAGGAGATCTACATCCAAACAGGTAGTGGTATTTTGGCATGATAT 310
QY 245 TCCCTGGTTGCTCTAGACATATGAAGAGCCCTCACACACAAGGTCGTGATCTCAGTCCC 304
Db 311 TCCCGGTTGCTCTAGACATTTGAAGACCT----- 342
QY 305 AAAGACCACCAAGAGCTCTCAAGGAGAGAGCAAAAGCCCAACAGCAACAGAGATAGTCACC 364
Db 343 -----CAACAAAAAGGACAAAGCAGCAGGCGCCCAAGACCGTCACC 382
QY 365 AGAAGTGACCGTTTCGATGAGGGTGATCTCATTTGCGATTCACCGCGTGTGCTTTCT 424
Db 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTCAGTGCACCGGTTTTGCATACT 442
QY 425 GGTCTTACAACGACCAACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAACAAC 484
Db 443 GGATGTACAACAATGAAGACACTCTCTGTTGTGGCGTTTCTTATTTGACACCAACAGCT 502
QY 485 ACGACAACCCAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAACACGGAGCAAG 544
Db 503 TCCAGNACCAGCTCGACCAGATGCTTAGGAGATTCTATCTTCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGAGCAAGAAAGCTTTACCATATAGCC 604
Db 563 AGTTTCTACAGTATCAGCCACAGAAAGCAGCAAG----- 595
QY 605 CATACAGCCGCAAGTCAAGCTAGCAAGAGAGCGTGAATTTAGCCCTCGAGGACAGC 664
Db 596 -----GAGGTACTC 604
QY 665 ACAGCCGAGAGAACGAGCAGGACGAAGAAGAAAGAAAGAGGTGGAAACATCTTTCAGCG 724
Db 605 AAAGCCAGAAAGGAAGCGTCAGCAAGAAGAAGAAAGAAAGAGGAGCAGCATATTGAGTG 664
QY 725 GCTTCAGCCCGAGTTCTTGGAAACAAAGCTTCCAGGTTGACGACAGACAGATAGTGCATA 784
Db 665 GCTTCGCCCGGAAATTTTGGAAACATGCGTTT-----GTGCTGACAGCAGGAGTAGTAGAA 721
QY 785 ACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGAGGAGCCATTTGTGACAGTGAAGGAGGCC 844
Db 722 AGCTCAAGGTGAGAACGAAGGAAGAGAGGGTGCCATTTGTGACAGTGAAGGAGGCTC 781

QY 665 ACAGCCGACAGACGAGGACAGAAAGAAACGAGGTGAAACATCTTCAGCG 724
Db |||||
QY 605 AAAGCCAGAAAGGAAAGCGTCAGCAAGAAAGAAACGAGGAGCGCATATTGAGTG 664
Db |||||
QY 725 GCTTCACGCGGAGTTCCTCGAAACAGCCTTCAGGTTGACGACAGACAGATAGTCAAA 784
Db |||||
QY 665 GCTTCGCGCGGAATCTTGAAACATCGCTC---GTCGTGGACAGCAGATAGTGAGAA 721
Db |||||
QY 785 ACCTAAGAGCGAGACCGAGAGTGAAGAGGAGGACCATTTGACAGTGAAGGAGGCGC 844
Db |||||
QY 722 AGCTACAAGGTGAGAACGAGAGGAGAGAGGAGGTCCTCATTTGACAGTGAAGAGGAGTC 781
Db |||||
QY 845 TCAGATCTTGAGCCCGCAGATAGAAAGAGAGCGTCCGACGAGAAAGAGGAATACGATGAG 904
Db |||||
QY 782 TCAGCGTGATAAGCCACCCAGCGAGAGAGCAGCAACAAAGACCGGAGGAGGAGAAAGC 841
Db |||||
QY 905 ATCAATATGATACGATGAAGAGATAGAAAGGCGTGCAGGGGAAAGCAGAGGAGGGGGA 964
Db |||||
QY 842 CAGATTGTGACGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
Db |||||
QY 965 ATGGTATTGAAGAGAGCATCTGACCGCAGAGTCTAAAGAAACATTTGGTAGAAACAGAT 1024
Db |||||
QY 887 ATGGCATTGACGAGACCATTTGCACATGAGACTTCGCCACACATTTGCCAGACTTCAT 946
Db |||||
QY 1025 CCCTGACATCTACACCCCTCAAGCTGCTTCACTCAAACTGCGCAAGCATCTCAACCTTC 1084
Db |||||
QY 947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCACACCGCTACCGCTCGACTTCG 1006
Db |||||
QY 1085 TAATACTTAGTGCTGTGGACCTAGTCTGAATATGMAATCTCTACAGGATGCAATCTG 1144
Db |||||
QY 1007 CAGCCCTCTGCTGGCTCAAACTCAGTGCCCGAGTTTGGATCACTCCGCAAGAAATGCTATGT 1066
Db |||||
QY 1145 TTGCTGCTCACTAAACACCAACGACACAGCATCATATATGATTTGAGGGGACGGGCTC 1204
Db |||||
QY 1067 TCGTGCCACACTACACCTGAACGCAACAGCATAATATACGCAATTTGAATGAGGGCAT 1126
Db |||||
QY 1205 ACCTGCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGTTCAGAGGGTC 1264
Db |||||
QY 1127 TGGTACAAGTGGTGAATTTGCAATGGTGCAGAGAGTGTGATGAGAGCTGCAAGAGGGAC 1186
Db |||||
QY 1265 ACCTGCTTGTGGTGCCACAGAACTTCGCGTCTGCGAAGTCCGAGAGGAGAACTCG 1324
Db |||||
QY 1187 AGGTGTTAATTTGGCCACAAACTTTGCGGTGCTGCAAGATCAGAGGCGCAACTTCG 1246
Db |||||
QY 1325 AATACGTGGCATTTCAAGACAGACTCAAGGCCAGCAGTACCACTCGCGGTGAAAACT 1384
Db |||||
QY 1247 AGTATGTTTCATTCAAGACCAATGATAGACCTCGATCGCAACCTTGAGGTGCAACT 1306
Db |||||
QY 1385 CGGTCAATAGAACCTTCGCGAGGAGGTGTTGCAAAATTCATATGCGCTCCAAAGGGAGC 1444
Db |||||
QY 1307 CATTGTTGAACGATTTGCGGAGGAGTGAATTCAGCAAACTTTTAACCTTAAGGAGGCAGC 1366
Db |||||
QY 1445 AGCAAGGAGCTTAAGAACAAACCCCTTCAAGTTCTTCAAGTTCTTCAAGCTCTCAGCAGT 1504
Db |||||
QY 1367 AGCCAGGAGGTCAGAAACAAACCCCTTCAAGTTCTTCAAGTTCTTCAAGCTCTCAGCAGT 1426
Db |||||
QY 1505 CTCGAGGGGCTGTGGCTTA 1523
Db |||||
QY 1427 AGAGGAGAGTTGTGGCTTA 1445
Db |||||

RESULT 12

US-10-757-667-13

; Sequence 13, Application US/10757667

; Publication No. US20040139504A1

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

;

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10757,667
; FILING DATE: 14-Jan-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-757-667-13

Query Match 34.0%; Score 517.4; DB 19; Length 1446;

Best Local Similarity 62.8%; Pred. No. 1.7e-154;

Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

QY 5 AGCAACGGAGAGAGAACGGTGCAGATTCCAGCGCTCAATCGCGAGAGACCTGACAATC 64
Db |||||
QY 71 AGCAGCCACAGCAAAACAGATGCCAGATCCAAAGCTCAATGCCCTAAACCCGGAATC 130
Db |||||
QY 65 GCATTGAATCAGAGCGCGTTACATTGAGACTTGGACCCCAACCAACAGGAGTTCGAT 124
Db |||||
QY 131 GTATAGAGTCAGAGGTGGCTTCATTGAGACATGGAACCCCTAACCAACAGCCATTCAGT 190
QY 125 GCGCGGGCTCGCCCTCTCTCGCTTAGTCTCGCGGCAACGCCCTTCGTAGGCCCTTCT 184
Db |||||
QY 191 GTGCGGGTTCGCCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTCGACAGACCTTCT 250
QY 185 ACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACCTTTGGTGGATAT 244
Db |||||
QY 251 ACACCAAGCTCCCGAGGAGATCTACCAACAAAGTAGTGGTATTTTGGCATGATAT 310
QY 245 TCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACACAGGTGTCGATCTCAGTCCC 304
Db |||||
QY 311 TCCCGGTTGCTCTAGACATTTGAAGAGCCT----- 342
QY 305 AAAGACCACCAAGAGCTCTCCAAGGAGAGAGCAAAAGCCCAACAGCAACGAGATAGTCACC 364
Db |||||
QY 343 -----CAACAAAAGGACAAAGCAGCAGCGGCCCAAGACCGTCAACC 382
QY 365 AGAAGGTGACCGTTTCGATGAGGTGATCTCATTTGCGAGTTCGCCACCGGTGTGCTTCT 424
Db |||||
QY 383 AGAAGATCTATCACTTCAGAGAGGAGTGTGATTTGATTCAGTGCAGTGCACACCGGTTTGCATAT 442
QY 425 GGCTCTACAAGCACACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAACAACA 484
Db |||||
QY 443 GGATGTACAACAATGAAGACACTCTCTGTGTTGCGGCTTCTCTTATTTGACACCAACAGCT 502

Qy	1025	CCCTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACTGCCAAGGATCTCAACCTTC	1084
Db	979		1038
Qy	1085	TAATACTTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCATTGT	1144
Db	1039		1098
Qy	1145	TTGTCGCTCACTACAACCAACGACACACAGCATCATATATATGAGGGGACGGGCTC	1204
Db	1099		1158
Qy	1205	ACGTGCAAGTCGTGGACAGCAACGGCNAACAGAGTGACGACGAGGAGCTTCAAGAGGGTC	1264
Db	1159		1218
Qy	1265	ACGTGCTTGTGGTGCCACAGAACTTCGCGCTCGTGGAAAGTCCACAGCGGAACTTCG	1324
Db	1219		1278
Qy	1325	AATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAAAC	1384
Db	1279		1338
Qy	1385	CCGTATAGATAACCTGCCGGAGGAGGTGGTGCAAAATTCATATGGCTCCAAAAGGGAGC	1444
Db	1339		1398
Qy	1445	AGCCAGGCAGCTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGT	1504
Db	1399		1458
Qy	1505	CTCCGAGGGCTGTGGCTTA	1523
Db	1459		1477

Search completed: August 24, 2005, 09:47:23
Job time : 931.248 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 86.3113 Seconds

(without alignments)
13658.075 Million cell updates/sec

Title:

Perfect score: 2691

Sequence: 1 CGGCAACACCGAGGAGGAA.....CTCCGAGGCGTGTGGCTTAA 1524

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124355_29200/app_query.fasta_1.4757
-DB=A_Geneseq_16Dec04 -OFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN_1_1_224 @runat_23082005_124355_29200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2673	99.3	510	6 ABUS2484	Abus2484 Peanut Ar
2	2673	99.3	510	7 ADG27544	Adg27544 Peanut al
3	2667	99.1	510	4 AAU04708	Aau04708 Anaphylac
4	2665	99.0	507	8 ADM12139	Adm12139 Arachis h
5	2663	99.0	510	2 AAY15246	Aay15246 Peanut al
6	2652	98.6	512	2 AAY40912	Aay40912 Ara h 3 a
7	2638	98.0	526	3 AAB33601	Aab33601 Modified
8	2638	98.0	526	4 AAU04711	Aau04711 Modified
9	2638	98.0	526	4 AAU05036	Aau05036 Modified
10	2615	97.2	507	8 ADO38357	Ado38357 Peanut al

11	1502	55.8	481	5 ABG71266	Abg71266 Glycine m
12	1502	55.8	481	7 ADH89253	Adh89253 G. max gl
13	1502	55.8	481	7 ADL90187	Adl90187 Soybean g
14	1502	55.8	481	8 ADG43988	Adg43988 G. max gl
15	1499.5	55.7	485	5 ABG71265	Abg71265 Glycine m
16	1499.5	55.7	485	7 ADH89247	Adh89247 G. max gl
17	1499.5	55.7	485	7 ADL90186	Adl90186 Soybean g
18	1499.5	55.7	485	8 ADG43982	Adg43982 G. max gl
19	1496	55.6	495	3 AAY80994	Aay80994 Soybean g
20	1496	55.6	495	5 ABG71264	Abg71264 Glycine m
21	1496	55.6	495	7 ADH89245	Adh89245 G. max gl
22	1496	55.6	495	7 ADL90168	Adl90168 Soybean g
23	1496	55.6	495	8 ADG43980	Adg43980 G. max gl
24	1496	55.6	511	7 ADL90190	Adl90190 Soybean g
25	1492.5	55.5	485	6 ABUS2502	Abus2502 Soybean g
26	1492.5	55.5	485	7 ADG27563	Adg27563 Soybean G
27	1488	55.3	484	2 AAY40949	Aay40949 Soybean g
28	1466	54.5	495	4 AAE10365	Aae10365 Soybean g
29	1034.5	38.4	457	8 ADO43062	Ado43062 Cashew nu
30	995.5	37.0	484	1 AAP71081	Aap71081 Sequence
31	986	36.6	517	5 ABG71267	Abg71267 Glycine m
32	986	36.6	517	7 ADL90189	Adl90189 Soybean g
33	973.5	36.2	562	7 ADH89249	Adh89249 G. max gl
34	973.5	36.2	562	7 ADL90188	Adl90188 Soybean g
35	973.5	36.2	562	8 ADG43984	Adg43984 G. max gl
36	946.5	35.2	562	5 ABG71268	Abg71268 Glycine m
37	941	35.0	561	1 AAP61363	Aap61363 Soybean g
38	926.5	34.4	516	8 ADH89251	Adh89251 G. max gl
39	926.5	34.4	516	8 ADG43986	Adg43986 G. max gl
40	914	34.0	185	2 AAY40951	Aay40951 Ara h 3 a
41	901.5	33.5	516	1 AAP61362	Aap61362 Soybean g
42	891	33.1	291	8 ADG60333	Ado60333 Cholester
43	845.5	31.4	499	7 ADC08275	Adc08275 Rice prot
44	841	31.3	499	7 ADC08135	Adc08135 Rice prot
45	841	31.3	499	7 ADC08279	Adc08279 Rice prot

ALIGNMENTS

RESULT 1

ABUS2484

ID ABUS2484 standard; protein; 510 AA.

AC ABUS2484;

DT 10-MAR-2003 (first entry)

DE Peanut Ara h3 protein sequence.

XX Peanut Ara h3 protein sequence.

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

KX anaphylactic food allergen; anti-allergenic; vaccine; wound healing.

XX Arachis hypogaea.

XX WO200274250-A2.

XX 26-SEP-2002.

XX 18-MAR-2002; 2002WO-US009108.

XX 16-MAR-2001; 2001US-0276822P.

XX 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;

PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

PI RabJohn PA, Shain DS, Stanley JS;

XX WPI; 2003-018765/01.

XX N-PSDB; ABX70612.

XX New modified anaphylactic food allergen, useful for preventing or

treating allergic reactions associated with e.g. anaphylactic allergens.

Claim 27; Fig 68B; 300pp; English.

The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IgE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding for or causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a peanut allergen (e.g. Ara hl, h2 or h3)

Sequence 510 AA;

Alignment Scores:
Pred. No.: 3.77e-236 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: Gaps: 0

US-10-728-323-3 (1-1524) x ABU52484 (1-510)

QY	1	CGCAGCAACCGAGAGAACGGTCCAGTTCACAGGCCCTCAATCGCGAGAGACCTGAC	60
DB	4	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	23
QY	61	AATCGGATTGAATCAGAGGCGGTACATTGACACTTGGAACTCCCAACACACAGGAGTTC	120
DB	24	AsnArgIleGluSerGluGlyGlyIleGluThrTrpAsnProAsnAsnGlnGluPhe	43
QY	121	GAATGGCGCGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGCAACGCCCTTCGTAGGCCT	180
DB	44	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	63
QY	181	TTCTACTCNAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGGTG	240
DB	64	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	83
QY	241	ATATTCCCTGGTGTCTTAGACACTATGAAGCCCTCACACAAAGGTCTGTCGATCTCAG	300
DB	84	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	103
QY	301	TCCAAAGACACCAAGAGCTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATGT	360
DB	104	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	123
QY	361	CACCAGAAGGTGACCGTTTCGATAGGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT	420
DB	124	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
QY	421	TTCTGGCTCTACCAACACCAACGACACTGATGTGTGTGTCTCTTACTGACACCAAC	480
DB	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
QY	481	AACAACGACACCAACGAGTTCATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACAGGAG	540
DB	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
QY	541	CAAGAGTTCTTAGGTTACCGCAACCAAGCAGCAACGAGCAAGAGAGCTTACCATAT	600

DB	184	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	203
QY	601	AGCCCATACAGCCCGCAAGAGTCAGCTAGCAAGAAGACGCGTAATTTAGCCCTCGAGGA	660
DB	204	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	223
QY	661	CAGCACAGCCGCGAGAGAACGACGACGACGAAGAAGAAACGAAGGTGGAAACATCTTC	720
DB	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	243
QY	721	ACGGCTTCACCGCGAGTTCCTGGAAACAGCCTTCCAGGTTCCAGACAGACAGATAGTG	780
DB	244	SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	263
QY	781	CAAAACCTTAAGAGCGGAGACCCAGAGTAGTGAAGAAGAGGAGCCATTGTGACGTGAGGGA	840
DB	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	283
QY	841	GGCCTCAGAAATCTTGAGCCCGACATAGAAAGACGCTGCCGACGAAGAGGAATACGAT	900
DB	284	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluTyrAsp	303
QY	901	GAAGATGAATATGAATCAAGATCAAGAGGATGAAGCGCTGGCAGGGAAGCAGAGGCAGG	960
DB	304	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	323
QY	961	GGGAATGGTATTGAAGAGAGACGATCTGCACCGCAAGTGCTAAAAAGAAACATTCGTAGAAC	1020
DB	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	343
QY	1021	AGATCCCTTGACATCTTACAACCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAAC	1080
DB	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	363
QY	1081	CTTCTAATACTTAGTGGCTTCGACCTAGTCTGATATATGGAATCTCTACAGGATGCA	1140
DB	364	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
QY	1141	TTGTTTGTGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGG	1200
DB	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
QY	1201	GCTCAGTGCACAGTCTGGACAGCAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG	1260
DB	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	423
QY	1261	GCTCAGTGTCTGTCGTCACAGAACTTCGCGCTCGTGGAAAGTCCCGACGAGCGAC	1320
DB	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
QY	1321	TTTCGAATACGTGGCATTCAAGACAGACTCAAGGCCCGACATAGCAACCTCCCGGTGAA	1380
DB	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
QY	1381	AACTCCGTATAGATAACTGCGGAGGAGTGTGCAAAATTCATAATATATGCGCTCAAAGG	1440
DB	464	AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuGlnArg	483
QY	1441	GAGCAGGACGACCTTAAGAACCAACCCCTTCAGTTCCTTCGTTCCACCGCTCTAG	1500
DB	484	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln	503
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	504	GlnSerProArgAlaValAla	510
RESULT 2			
ID	ADG27544		
XX	ADG27544	standard; protein; 510 AA.	
AC	ADG27544;		
XX	DT	26-FEB-2004 (first entry)	

XX Peanut allergen Ara h3.
 XX
 XX Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
 KW antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.
 XX
 XX Arachis hypogaea.
 XX
 XX US2003202980-A1.
 XX
 XX 30-OCT-2003.
 XX
 XX 18-MAR-2002; 2002US-00100303.
 XX
 XX 29-DEC-1995; 95US-0009455P.
 PR 23-SEP-1996; 96US-00717933.
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 29-JUN-1998; 98US-00106872.
 PR 27-AUG-1998; 98US-00141220.
 PR 13-NOV-1998; 98US-00191593.
 PR 29-JAN-1999; 99US-00240557.
 PR 29-JAN-1999; 99US-00241101.
 PR 11-FEB-1999; 99US-00248673.
 PR 11-FEB-1999; 99US-00248674.
 PR 02-MAR-1999; 99US-0122450P.
 PR 02-MAR-1999; 99US-0122452P.
 PR 02-MAR-1999; 99US-0122560P.
 PR 02-MAR-1999; 99US-0122565P.
 PR 11-MAR-1999; 99US-00267719.
 PR 28-JAN-2000; 2000US-00494096.
 PR 16-MAR-2001; 2001US-0276822P.
 XX
 XX (CAPL/) CAPLAN M J.
 PA (SOSI/) SOSIN H B.
 PA (SAMP/) SAMPSON H.
 PA (BANN/) BANNON G A.
 PA (BURK/) BURKS A W.
 PA (COCK/) COCKRELL G.
 PA (COMP/) COMPADRE C M.
 PA (CONV/) CONNAUGHTON C.
 PA (HELM/) HELM R M.
 PA (KING/) KING N E.
 PA (KOPP/) KOPPER R A.
 PA (MALE/) MALEKI S J.
 PA (RABJ/) RABJOHN P A.
 PA (SHIN/) SHIN D S.
 PA (STAN/) STANLEY J S.
 XX
 XX Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabbjohn PA, Shin DS, Stanley JS;
 XX
 XX WPI; 2003-875632/81.
 DR N-PSDB; ADG27543.
 XX
 XX New modified anaphylactic food allergen comprising a cysteine residue
 PT which has been modified so that it cannot participate in the disulfide
 PT bond, useful for treating allergic reactions or wounds.
 XX
 XX Example 16; SEQ ID NO 89; 194pp; English.
 XX
 XX The invention relates to a modified anaphylactic food allergen whose
 CC amino acid sequence is substantially identical to that of a natural
 CC anaphylactic food allergen. The natural anaphylactic food allergen
 CC includes at least one cysteine residue that participates in a disulphide
 CC bond when the natural anaphylactic food allergen is in its native
 CC conformation, except that the cysteine residue has been modified so that
 CC it cannot participate in the disulphide bond. Also included are a method
 CC of making a modified anaphylactic food allergen, a nucleotide molecule

CC encoding a modified anaphylactic food allergen defined above, a
 CC nucleotide molecule for causing a site specific mutation in a gene
 CC encoding a natural anaphylactic food allergen, a transgenic plant or
 CC animal expressing a modified anaphylactic food allergen defined above, a
 CC method of treating an individual by reducing the clinical response to a
 CC natural anaphylactic food allergen by administering a modified
 CC anaphylactic food allergen and an isolated fragment of peanut allergen
 CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
 CC ADG27465. About 10-17% of the amino acids have been modified in at least
 CC one IGE epitope or all the IGE epitopes recognised when the natural
 CC anaphylactic food allergen is contacted with serum IGE from individual(s)
 CC allergic to the natural anaphylactic food allergen. The invention
 CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
 CC cDNAs), Soybean Glycinin A2B1a and IGE-binding epitopes of the English
 CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
 CC used for treating allergic reactions or wounds. The present sequence
 CC represents a Peanut allergen of the invention (or its fragment).
 XX
 XX SQ Sequence 510 AA;
 Alignment Scores:
 Pred. No.: 3,77e-236 Length: 510
 Score: 2673.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.33% Indels: 0
 DB: 7 Gaps: 0
 US-10-728-323-3 (1-1524) x ADG27544 (1-510)
 QY 1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTCACGGCGCTCAATGGCGAGACCTGAC 60
 DB 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
 QY 61 AATCGCATTGAATCAGAGGCGGTATACATTGACACTTGGAACTCCCAACCCCAACACAGGAGTTC 120
 DB 24 AsnArgIleGluSerGluGlyIleGluThrTyrAsnProAsnGlnGluPhe 43
 QY 121 GAATGCGCGGCGTCCCGCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGGCCT 180
 DB 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
 QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAGAGGAGGATCTTTGGTTG 240
 DB 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
 QY 241 ATATTCCTCTGTTGTCTTAGACACTATGAAGAGCCTCACACAAAGTCGTCTGATCTCAG 300
 DB 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
 QY 301 TCCCAAGACCAACCAAGACGTCTTCCAGGAGAGAGCAACCAAGCCACAGCAACGAGATAGT 360
 DB 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
 QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCTATTCAGTTCCTCCACCGGTGTGCT 420
 DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 QY 421 TTCTGCTCTTACAAACGACCAACGACACTGATGTTGTTGTTGTTCTTCTTACTGACCAAC 480
 DB 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
 QY 481 AACACGACACCAACGAGCTTGTATCAGTTCCCGAGGAGATTCAATTGGCTGGGAACACGGAG 540
 DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183
 QY 541 CAAGATCTTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGAGCAAGACCTTACCATAT 600
 DB 184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
 QY 601 AGCCCATACAGCCCGCAAGTAGTCCAGCTTACAGAAAGAGCGTGAATTTAGCCCTCCGAGGA 660
 DB 204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223

QY 661 CAGCAGCCGCGAGAACGAGCAGGACAAAGAAAGAAACGAAAGTGGAAACATCTTC 720
 Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
 QY 721 AGCGGCTTACCGCGAGTTCCTGGAAAGACCTTCAGGTTGACGACAGACAGATAGTG 780
 Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal 263
 QY 781 CAAACCTAGAGCGCAGCCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
 Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 283
 QY 841 GGCTCAGATCTTGGACCCAGATAGAAAGAGACGTCGCCACGAAGAAGAAATACGAT 900
 Db 284 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 303
 QY 901 GAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGAAGCAGGAGG 960
 Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
 QY 961 GSGAATGGTATTGAGAGAGCATCTGCACCGCAAGTGTAAAGAACATTGGTAGAAC 1020
 Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleGlyAsnIleGlyArgAsn 343
 QY 1021 AGATCCCTCAGCATCTCAACACCTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAAC 1080
 Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuIleThrAlaAsnAspLeuAsn 363
 QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGAATATGAAATCTTACAGGAATGCA 1140
 Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
 QY 1141 TTCTTTGTGCTCAGTACACCAACGACGACGATCATATATCGATTGAGGACCG 1200
 Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
 QY 1201 GCTCAGCTGCAAGTCTGGACGACCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
 Db 404 AlaHisValGlnValAlaAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
 QY 1261 GGTCAAGTGTGTTGGTGCACAGAACTTCGCGCTGCTGGAAAGTCCACAGCAGGAAC 1320
 Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
 QY 1321 TTCGAATACGTGCTTCAAGACAGACTCAAGGCCAGCATGCCAACTCGCGGTGAA 1380
 Db 444 PheGluTyrValAlaPheIleThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
 QY 1381 AACTCCGTATAGATAACCTGCGGAGGAGGTGGTTGCCAAATTCATATGCGCTCCAAAGG 1440
 Db 464 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuGlnArg 483
 QY 1441 GAGCAGCAGGAGGAGCTTAAGAACACACCCCTTCAAGTTCCTGTTCCACGCTCAG 1500
 Db 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProSerGln 503
 QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
 Db 504 GlnSerProArgAlaValAla 510

RESULT 3
 AAU04708
 ID AAU04708 standard; protein; 510 AA.
 XX AC AAU04708;
 XX XX
 DT 23-OCT-2001 (first entry)
 XX Anaphylactic antigen Ara h 3.
 DE Anaphylactic antigen Ara h 3.
 XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse.

XX OS Mus sp.
 XX PN WO200140264-A2.
 XX PD 07-JUN-2001.
 XX PF 06-DEC-2000; 2000WO-US033124.
 XX PR 06-DEC-1999; 99US-00455294.
 XX PR 23-JUN-2000; 2000US-021376SP.
 XX PR 27-SEP-2000; 2000US-0235797P.
 XX {PANA-} PANACEA PHARM LLC.
 PA {UYAR-} UNIV ARKANSAS.
 PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
 XX Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
 PI WPI; 2001-381378/40.
 XX DR N-PSDB; AAS08540.
 XX Antigenic fragments useful for reducing anaphylactic risk and reducing
 PT the severity and/or number of allergic symptoms in individuals sensitive
 PT to antigens, have reduced ability to bind Immunoglobulin E.
 XX Claim 8; Fig 11; 100pp; English.
 XX The sequence represents the amino acid sequence of anaphylactic antigen
 CC Ara h 3. Ara h 3 is an anaphylactic antigen (A), which was used to design
 CC antigenic peptides having a reduced ability to bind immunoglobulin E
 CC (IgE) as compared with the intact (A), or having a sequence substantially
 CC identical to a portion of sequence of an antigen that includes at least
 CC one IgE binding site, where at least one IgE binding site of the peptide
 CC is altered. The antigenic peptides are used in a composition which is
 CC useful for reducing risk or severity of allergic reaction to an antigen.
 CC This is done by identifying an individual at risk of allergic reaction to
 CC an antigen by identifying prior display of allergic symptoms when exposed
 CC to the antigen, or a familial relationship with an individual who
 CC previously displayed allergic symptoms when exposed to the antigen.
 CC Following this an antigen-specific IgE present on one or more mast cells
 CC or basophils in the individual's serum is identified. The individual is
 CC then contacted with a peptide corresponding to a portion of the antigen,
 CC which is selected, formulated, and delivered so that binding of the
 CC peptide to antigen-specific IgE is reduced as compared with IgE binding
 CC of intact antigen. The composition is also useful for treating and
 CC preventing allergic reactions
 XX SQ Sequence 510 AA;

Alignment Scores:
 Pred. No.: 1,34e-235 Length: 510
 Score: 2667.00 Matches: 506
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.11% Indels: 0
 DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x AAU04708 (1-510)
 QY 1 CGGCAGCAACCGAGGAGAAACGCGTCCAGCGCTCAATGCCGAGACCTGAC 60
 Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
 QY 61 AATCCGATTGAATCAGAGCGGTTACATTGAGATTGGAAACCCCAACACGAGGTTTC 120
 Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
 QY 121 GAATGCGCGCGGTGCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTCTGTAGGCT 180
 Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
 QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACATTTGGGTG 240

Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
QY 301 TCCCAAGACACCAAGACGCTCCCAAGGAGAGACCAAGGCAACAGCAAGAGATAGT 360
Db 104 SerGlnArgProProArgArgLeuGlnGlyAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAAGGTGCACCGCTTTTCAGATGAGGGTGATCTCATTCAGTCTCCACCGGTGTGCT 420
Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTTCTGCTCTACACGACACGACACTGATGTGTGCTGCTGTTCTCTTACTGACCAAC 480
Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACACGACAAACGCTTGATCAGTTCCCGCAGGAGATTCATTTGGCTGGGAACAGGAG 540
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTCTTAAAGTTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db 184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY 601 AGCCCATACAGCCCGCAAGCTAGCCTTAGACAAGAAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
QY 661 CAGCAGCCGCGAGAGAACGAGCAGACGACGAAGAGAAAGCAAGAGTGAACATCTTC 720
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 243
QY 721 AGCGGCTTCACGCGCGAGTCTTCGGAACAAGCCTTCAGGTTCCACGACAGACAGATAGT 780
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAAACTTAAGAGCGGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGGGGGA 840
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
QY 841 GGCTCAGAACTCTTGAGCCGAGATAGAAAGACGTCGCCGACGAAGAGGAATACGAT 900
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
QY 901 GRAGATCAATATCAATACGATGAAGGATAGAGGCGTGCAGGGGAGCAGGAGCAGG 960
Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyThr 323
QY 961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAAAGAACATTGGTAGAAAC 1020
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY 1021 AGATCCCTCAGATCTACAAACCTCAAGCTGGTTCATCTCAAACTGCCAACGATCTCAAC 1080
Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATCTAGGTGGCTTGACCTAGTGTGTAATATATGGAATCTCTACAGGAATGCA 1140
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
QY 1141 TTGTTTGTGCTCACTACACCAACGACGACACGATCATATATCGATTGAGGGGACGG 1200
Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY 1201 GCTCAGTGAAGTCTGTGGACAGCAACGCGCAACAGAGTGTACGAGGAGGCTTCAAGAG 1260
Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY 1261 GGTCACTGCTGTGTCACAGAACTTCGCGTTCGCTGGAAGTCCAGGCGAGAAC 1320

Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380
Db 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCGTCTATAGATTAACCTGCCGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAG 1440
Db 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
QY 1441 GAGCAGCAAGGAGGCTTAAGAACAAACCCCTTCAAGTTCTTCTGTTCCACCGTCTCAG 1500
Db 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
QY 1501 CAGTCTCCAGGCGCTGTGCT 1521
Db 504 GlnSerProArgAlaValAla 510
RESULT 4
ADM12139
ID ADM12139 standard; protein; 507 AA.
XX
AC ADM12139;
XX
DT 20-MAY-2004 (first entry)
XX
DE Arachis hypogaea 3 (Ara h3) protein.
XX
KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.
XX
OS Arachis hypogaea.
XX
PN US2003235594-A1.
XX
PD 25-DEC-2003.
XX
PF 17-SEP-2002; 2002US-00245871.
XX
PR 14-SEP-1999; 99US-00396813.
PR 17-JUL-2002; 2002US-00197000.
XX
PA (ANTI-) ANTIGEN EXPRESS INC.
XX
PI Humphreys R, Xu M;
XX
PS WPI; 2004-070554/07.
XX
CC Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
CC for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
CC erythematosus and diabetes mellitus.
XX
PS Example 3; Page 21; 87pp; English.
XX
CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus

CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 CC represents a mammalian II key related protein of the invention.
 XX
 SQ Sequence 507 AA;
 Alignment Scores:
 Pred. No.: 2,04e-235 Length: 507
 Score: 2665.00 Matches: 506
 Percent Similarity: 99.80% Conservativeness: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.03% Indels: 0
 DB: Gaps: 0
 US-10-728-323-3 (1-1524) x ADM12139 (1-507)
 QY 1 CGCAGCAACCGAGAGAACGGTCCAGTTCACGGCCTCAATGCGGAGAGACTGAC 60
 Db
 QY 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
 Db
 QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGAACCCCAACACAGGAGTTC 120
 Db
 QY 21 AsnArgIleGluSerGluGlyGlyTyrlleGluThrIrpAsnProAsnAsnGlnGluPhe 40
 QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
 Db
 QY 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 60
 QY 181 TTCTACTCCNATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTG 240
 Db
 QY 61 PheTySerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrllePheGlyLeu 80
 QY 241 ATATTCCCTGTTGTCTAGACACTATGAAGACCTCACACAAAGTCTGTCGATCTCAG 300
 Db
 QY 81 IlePheProGlyCysProArgHisTyrlleGluProHisThrGlnGlyArgArgSerGln 100
 QY 301 TCCCAAGACCAACAGAGCTCTCAAGGAGAGAACAAAGCCCAACAGCAACGAGATAGT 360
 Db
 QY 101 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
 QY 361 CACGAGAGGTGACGGTTTCGATGAGGTGATCTCATTCGAGTTCGCCACCGGTGTGCT 420
 Db
 QY 121 HisGlnLysValHisArgThrAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
 QY 421 TTCTGGCTCTACACACACACACTGATGTTGTTGTTCTTCTACTGACACCAAC 480
 Db
 QY 141 PheTrpLeuTyrlleAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
 QY 481 AACACGACCAACGAGTTCATGATCTCCCGAGGAGATTCAAATTTGGCTCGGCAACGGAG 540
 Db
 QY 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
 QY 541 CAAGAGTCTTAAGTACCGCAACAAAGCAGACAAAGCAGAGAGAGAGTTCACCATAT 600
 Db
 QY 181 GlnGluPheLeuArgTyrlleGlnGlnSerArgGlnSerArgArgArgSerLeuProTyrl 200
 QY 601 AGCCCATACAGCCGCAAGTCTAGCTAGCAAGAAGAGCGTCAATTTAGCCCTCGAGGA 660
 Db
 QY 201 SerProTyrlSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 220
 QY 661 CAGCAGCCGCGAGAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
 Db
 QY 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 240
 QY 721 AGCGGCTTCAGCGGAGTTCCTGGACACAGCCCTTCAGGTTGACACAGACAGATAGTG 780
 Db
 QY 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
 QY 781 CAAACCTTAAGAGCGAGACCGAGAGTGAAGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 840
 Db
 QY 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
 QY 841 GGCCTCAGAACTTTGAGCCCGAGATGAAGAAGAGAGAGTGTCCCGACGAGAGGAATACGAT 900

Db 281 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluGluGluGlu 300
 QY 901 GAAGATGAATATGAATACGATGAAGAGGATGAAGCGCTGCGAGGAGGAGGAGGAGGAGG 960
 Db 301 GluAspGluTyrlleGluTyrlleGluAspGluGluAspArgArgGlyArgGlyArgGly 320
 QY 961 GCGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAAAGAACAACTTGGTAGAAAC 1020
 Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAlaLysAlaGlnGlyArg 340
 QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAAC 1080
 Db 341 ArgSerProAspIleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeu 360
 QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGATGCA 1140
 Db 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArg 380
 QY 1141 TTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
 Db 381 LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArg 400
 QY 1201 GCTCAGTCCAAAGTCTGTGACAGCAACGGCAACAGAGTGTACGAGAGGCTTCAAGAG 1260
 Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluGlnGlu 420
 QY 1261 GGTCAAGTCTGTTGGTCCACAGAACTTCGCGCTGCGTGGAAAGTCCCGAGGAGGAG 1320
 Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGlu 440
 QY 1321 TTTCGAATACGTGGCATTCAAAGACAGACTCAAGGCGGACAGCATAGCCAACTCCCGGTGAA 1380
 Db 441 PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGly 460
 QY 1381 AACTCGTCTATAGATAAATCTGCGGAGGAGGTGTTGCAAAATTCATATGCGCTCCAAAGG 1440
 Db 461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrlleGlyLeuGln 480
 QY 1441 GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTCAAGTCTTCGTTCCACCGTCTCAG 1500
 Db 481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProSerGln 500
 QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
 Db 501 GlnSerProArgAlaValAla 507
 RESULT 5
 ID AAY15246
 XX AAY15246 standard; protein; 510 AA.
 AC AAY15246;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-NOV-1999 (first entry)
 XX
 DE Peanut allergen, Ara h 3, amino acid sequence.
 XX
 KW allergy; immune response; transgenic; allergen; epitope;
 KW immunoglobulin E; Ig E; binding site; peanut.
 XX
 OS Arachis hypogaea.
 XX
 PN WO9938978-A1.
 XX
 PD 05-AUG-1999.
 XX
 PF 29-JAN-1999; 99WO-US002031.
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.

PR 27-AUG-1998; 98US-00141220.
 XX (UYAR-) UNIV ARKANSAS.
 PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) SOSIN H.
 XX Sosin H, Bannon GA, Burks AW, Sampson HA;
 XX WPI; 1999-479189/40.
 DR N-PSDB; AA206384.
 XX Modified allergen with reduced IgE binding, useful for treating e.g.
 XX allergies.
 XX Disclosure; Page 39-40; 46pp; English.
 XX This is the amino acid sequence of the Ara h 3 protein from *Arachis*
 CC hypogaea. The Ara h 3 protein has 4 IgE (Immunoglobulin E) binding
 CC epitopes, one of which is immunodominant (AAV15281). Note: The first
 CC three amino acids of this sequence are not encoded by the related
 CC nucleotide sequence. By modifying the IgE binding sites the ability of
 CC the allergen to provoke an immune response is downregulated. The epitopes
 CC of the IgE binding sites can therefore be modified in genetically
 CC engineered plants and animals to elicit less of an allergic response.
 CC (Updated on 17-Oct-2003 to standardise OS field)
 XX Sequence 510 AA;
 SQ

Alignment Scores:
 Pred. No.: 3.13e-235 Length: 510
 Score: 2663.00 Matches: 506
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 98.96% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-3 (1-1524) x AAV15246 (1-510)

QY 1 CGGCACAAACCGGAGAGAACGGTCCAGTTCAGCGGCTCAATCGCAGAGACCTGAC 60
 Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
 QY 61 AATCGCATTAATCAGAGGCGGTACATTGACATTGGAAACCCCAACACAGGAGTTC 120
 Db 24 AsnArgIleGluSerGluGlyGlyIleGluThrTrpAsnProAsnGlnGluPhe 43
 QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCCT 180
 Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
 QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTG 240
 Db 64 PheTy-SerAsnAlaProGlnGlnIlePheIleGlnGlnGlyArgGlyTyPheGlyLeu 83
 QY 241 ATATTCCCTGGTTGCTTAGACTATGAAGACCTCACACAAAGTGTGATCTCAG 300
 Db 84 IlePheProGlyCysProArgHisTyGluGluProHisThrGlnGlyArgArgSerGln 103
 QY 301 TCCCAAGACCAACAGAGCTCTCCAAAGGAGAGACCAAGCCCAACAGCAACGAGATGT 360
 Db 104 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
 QY 361 CACCAGAAGGTGACCGTTTCCGATGAGGGGTGATCTCATTTGCAGTTCCTCCACCGGTGTGCT 420
 Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 QY 421 TTCTGCTCTACCAACACGACGACTGATGTGTGTGTTGTTCTTCTTACTGACCAAC 480
 Db 144 PheTrpLeuTy-SerAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
 QY 481 AACCAACGACCAACCGTTCATCAGTTCCTCCAGGAGATTCAATTTGCTGGGAACCGGAG 540
 Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183

QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAACGACAGACAAAGCAGACGAAGAGCTTACCATAT 600
 Db 184 GlnGluPheLeuArgTyGlnGlnGlnSerArgGlnSerArgArgSerLeuProTy 203
 QY 601 AGCCCATACAGCCCGCAAGTTCAGCTACGAAGAGAGCGTCAATTTAGCCCTCCAGGA 660
 Db 204 SerProTy-SerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
 QY 661 CAGCACAGCCGACAGAACGACGACGACGAAGAAAGAAACGAGTGGAAACATCTTC 720
 Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
 QY 721 AGCGCTTCACGCCGAGTTCTCTGGAAACAAGCTTCCAGGTTGACACAGACAGATAGTG 780
 Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
 QY 781 CAAACCTTAAGAGGCGAGACCCAGAGTAGTGAAGAGGAGGCCATTCTGCAGTGAGGGA 840
 Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
 QY 841 GGCCTCAGAAATCTTGAGCCCGACATAGAAAGACAGCGTCCGACGAAAGAGGAATACGAT 900
 Db 284 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyAsp 303
 QY 901 GAAGATGAATATGATACGATGAAGAGTAGAGCGCTGGCAGGGGAAAGCAGAGCGAGG 960
 Db 304 GluAspGluTy-GluTyAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
 QY 961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAGAAACATTCGTAGAAAC 1020
 Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
 QY 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCATCTCAAACTCCCAACGATCTCAAC 1080
 Db 344 ArgSerProAspIleTyAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
 QY 1081 CTTCTAATCTTAGTGGCTTCGACCTAGTCTGAATATATGAAATCTCTACAGGAATGCA 1140
 Db 364 LeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTy-GlyAsnLeuTyArgAsnAla 383
 QY 1141 TTGTTTGTGCTCCTACCAACCAACGACACGATCATATATCGATTGAGGGGACGG 1200
 Db 384 LeuPheValAlaHisTyAsnThrAsnAlaHisSerIleIleTyArgLeuArgGlyArg 403
 QY 1201 GCTCAGTGCACAGTCTGGACACCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
 Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyAspGluGluGlnGlu 423
 QY 1261 GGTACGCTTGTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAAC 1320
 Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaLysSerGlnSerGluAsn 443
 QY 1321 TTTCGAATAGTGGCATTCGAAGCAGACTCAAGCCCGCAGCATAGCCAACTCCCGGTGAA 1380
 Db 444 PheGluTyValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
 QY 1381 AACTCGCTCATAGATAACCTCGCGGAGGAGTGTGCAATTCATATGGCTCCAAAGG 1440
 Db 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyGlyLeuGlnArg 483
 QY 1441 GAGCAGGCAAGCGACTTAAGAAACAAACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
 Db 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProProSerGln 503
 QY 1501 CAGTCTCCGAGCGCTGTGGCT 1521
 Db 504 GlnSerProArgAlaValAla 510

RESULT 6
 ID AAV40912
 XX AAY40912 standard; protein; 512 AA.

Db 404 GlyArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeu 423
 Qy 1255 CAAGAGGTCACGTGCTTGTGTGTCACAGAACTTCGCCGTCGTGGAAAGTCCAGAGC 1314
 Db 424 GlnGluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyIysSerGlnSer 443
 Qy 1315 GAGAACTTCGAATACGTGGCATTCACAGACACTCAAGGCCAGCATAGCCCACTCGCC 1374
 Db 444 GlnAsnPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAla 463
 Qy 1375 GGTGAAACTCCGTCATAGATAACTCCCGAGGAGGTGTTGCMAATTCATATGGCCTC 1434
 Db 464 GlyGluAsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeu 483
 Qy 1435 CAAGAGGAGCAGGCAAGCGAGCTTAAGAACAAACCCCTTCAGTTCTTCGTTCCACCG 1494
 Db 484 GlnArgGluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProPro 503
 Qy 1495 TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
 Db 504 SerGlnGlnSerProArgAlaValAla 512

RESULT 7

AAB33601

ID AAB33601 standard; protein; 526 AA.

XX

AC AAB33601;

XX

XX 12-SEP-2003 (revised)

DT 22-JAN-2001 (first entry)

XX

DE Modified Ara h 3 amino acid sequence.

XX

KW Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;
 Kw Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;
 Kw food antigen; sensitising; immune response; anti-allergic.

XX

OS Arachis hypogaea.

XX

PN WO200051647-A2.

XX

PD 08-SEP-2000.

XX

PF 03-MAR-2000; 2000WO-US005655.

XX

PR 03-MAR-1999; 99US-0122960P.

XX

PR 06-DEC-1999; 99US-00455294.

XX

PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX

PA Sampson HA;

XX

PI WPI; 2000-611341/58.

XX

PT Non-human animal sensitized to an antigen, useful as an animal model for
 PT studying allergic reactions to allergens, such as those in food and in
 PT the environment.

XX

PS Example 6; Fig 17C; 124pp; English.

XX

CC The present invention describes an animal model which can be used for
 CC studying allergic reactions to allergens. The animal is sensitised to a
 CC selected antigen by administering the antigen itself or a nucleic acid
 CC encoding the antigen, where preferably the antigen is an anaphylactic
 CC antigen. The sensitised animal can then be used to screen for compounds
 CC which may help to prevent, ameliorate, or cure allergic conditions in
 CC humans. The animal model can be used for studying allergic reactions to
 CC allergens, such as those in food (peanuts, fruit, berry, nut, bean, milk,
 CC dairy products), or in the environment (weed pollen, grass pollen, tree
 CC pollen, mite, animal, animal dander, fungal, and insect antigens).
 CC AAB33478 to AAB33601 represent sequences which are used in examples from
 CC the present invention to specifically examine the peanut allergy, and the

CC peanut anaphylactic antigens Ara h 1 to 3. (Updated on 12-SEP-2003 to
 CC standardise OS field)

XX

SQ Sequence 526 AA;

XX

Alignment Scores:

Pred. No.: 6,27e-233 Length: 526
 Score: 2638.00 Matches: 502
 Percent Similarity: 99.01% Conservative: 0
 Best Local Similarity: 99.01% Mismatches: 5
 Query Match: 98.03% Indels: 0
 DB: 3 Gaps: 0

US-10-728-323-3 (1-1524) x AAB33601 (1-526)

Qy 1 CGGCACAAACCGGAGGAGAACGCGTCCAGTTCACGCGCTCAATGCGCAGACCTTGAC 60

Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24

Qy 61 AATCGCATTCGAATCAGAGGCGGTTCATTGAGACTTGGNAACCCCAACACGAGGATTC 120

Db 25 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnAlaAsnGlnGluPhe 44

Qy 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180

Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgPro 64

Qy 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGATCTTTGGGTTG 240

Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84

Qy 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300

Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104

Qy 301 TCCCAAGACACCAAGAGCTCTCCAAGGAGAGACCAAGCAACAGCAACGAGATAGT 360

Db 105 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124

Qy 361 CACCAGAGGTGCACCGTTTTCGATGAGGTGATCTCATTCGAGTTCACCGCGTGTGCT 420

Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144

Qy 421 TTCTGCTCTCAACAGCACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480

Db 145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164

Qy 481 AACACGACAAACGAGCTTCATCAGTTCCCGAGGAGATTCAATTTGGCTGGACACGGAG 540

Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 184

Qy 541 CAAGAGTCTTAAAGGTACCAACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600

Db 185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204

Qy 601 AGCCCATACAGCCCGCAAGAGTCCAGCTAGCAAGAAAGAGCGTGAATTTAGCCCTCAGGA 660

Db 205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224

Qy 661 CAGCAGCCGCGAGAGAACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720

Db 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluGlu 244

Qy 721 AGCGGCTTCACCGCGAGTTCCTCGAACCAAGCCTTCAGGTTTCACGACGACGACGATAGTG 780

Db 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264

Qy 781 CAAACCTTAAGAGCGGAGACCCGAGAGTGAAGAGAGGAGGCCATTGTGACGTAGGGGA 840

Db 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGlu 284

Qy 841 GGCCTCAGAACTTTGAGCCCGACAGACGAGTCCGCGACGACGACGACGACGACGACGAT 900

Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATGAATATATACGATGAAGAGGATAGAAGCGGTGGCAGGCGAAGCAGAGCAGG 960
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlySerArgGlyArg 324
QY 961 GGAATGGTATTGAAGAGAGCATCTCGACCGCAAGTGTCTAAAGAACATTGGTAGAAGC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTGACATCTACACCTCAAGCTGCTTCACTCAAACTGCCAACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATACTTAGTGGGTGGTGGACCTAGTGCCTGAATATGGAATCTTACAGGAATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
QY 1141 TTGTTTGTGCTCACTACCAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
QY 1201 GCTCAGTGCAGTCTGTGACACCAACGACAGAGTGTACGACGAGGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluLeuGlnGlu 424
QY 1261 GGTCACTGCTGCTGTGTCGCACAGAACTTCGCCGTGCTGGAAGTCCACAGACGAGAAC 1320
Db 425 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTCGAATAGTGGCATTCAGACAGACTCAAGCCAGCATAGCCAACTCCGCCGTGAA 1380
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTATAGATTAACCTGCGGAGGAGGTGTTGCAATTCATATGCGCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484
QY 1441 GAGCAGGCAAGGAGGAGTAAAGAACAAACACCCCTTCAAGTTCTTCCACCGCTCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProSerGln 504
QY 1501 CAGTCTCCGAGGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

RESULT 8
AAU04711
AC AAU04711 standard; protein; 526 AA.
XX AAU04711;
DT 23-OCT-2001 (first entry)
XX
DE Modified anaphylactic antigen Ara h 3.
XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO200140264-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033124.
XX
PR 06-DEC-1999; 99US-00455294.
PR 23-JUN-2000; 2000US-0213765P.
PR 27-SEP-2000; 2000US-0235797P.
XX
PA (PANA-) PANACEA PHARM LLC.

(UYAR-) UNIV ARKANSAS.
(MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
WPI; 2001-381378/40.
Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E.
Disclosure; Fig 12; 100pp; English.
The sequence represents the amino acid sequence of modified anaphylactic antigen Ara h 3, which has been altered to disrupt immunoglobulin E (IgE) binding sites. Ara h 3 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind IgE as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who previously displayed allergic symptoms when exposed to the antigen. Following this an antigen-specific IgE present on one or more mast cells or basophils in the individual's serum is identified. The individual is then contacted with a peptide corresponding to a portion of the antigen, which is selected, formulated, and delivered so that binding of the peptide to antigen-specific IgE is reduced as compared with IgE binding of intact antigen. The composition is also useful for treating and preventing allergic reactions

SQ Sequence 526 AA;
Alignment Scores:
Pred. No.: 6,278-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 5
Query Match: 98.03% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x AAU04711 (1-526)
QY 1 CGGCAGCAACCGGAGGAGAACGCGTCCAGCGCTCAATGCGCAGAGACTGAC 60
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGCATTTGAATCAGAGGGCGTTACATTGAGACTTGGAAACCCCAACACGAGAGTTC 120
Db 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
QY 121 GAATCGCGCGCGTCCGCTCTCTCGTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 64
QY 181 TTCTACTCCATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATCTTTCGGTTG 240
Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCCTGGTGTGCTAGACACTATGAAGAGCTTCACACAAGGCTCGTCGATCTCAG 300
Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
QY 301 TCCCAAGACCAACCAAGACGCTCCAGAGAGAAGACCAAGCAACAGCAAGATAGT 360
Db 105 SerGlnArgProProArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
QY 361 CACCAGAGGTGCACCGTTTCGATCAGGTGATCTCATTGCAAGTTCACACCGGTGTCT 420
Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144

US-10-728-323-3 (1-1524) x AAU05036 (1-526)			
QY	1	CGGCACCAACGGAGAGACGGTCCAGTTCACGGCTCAATCGCGAGACCTGAC	60
Db	5	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	24
QY	61	AATCGCATTAATACAGAGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC	120
Db	25	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTTPAsnAlaAsnAsnGlnGluPhe	44
QY	121	GAATCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGCAACGCCCTTCGTAGCCT	180
Db	45	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	64
QY	181	TTCTACTCCAATGCTCCCGAGGAGTCTCATCCAGCAAGGAGGGATCTTTGGGTTG	240
Db	65	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	84
QY	241	ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCCTCACACAAGGTGCTCGATCTCAG	300
Db	85	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	104
QY	301	TCCCAAGACCACAAGACTCTCCAGGAGAAAGACCAAGCCAAACAGCAACGAGATAGT	360
Db	105	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	124
QY	361	CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCACCGGTGTGCT	420
Db	125	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	144
QY	421	TTCTGGCTCTACACGACCACGACACTGATGTTGTTGCTGTTTCTTACTGACACCAAC	480
Db	145	PheTyrLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	164
QY	481	AACAACGACACCGCTTGATGATTCATGTTCCCGAGAGATTCAATTTGCTGGGAACACGGAG	540
Db	165	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	184
QY	541	CAAGAGTCTTAAGTACCGACCAACAAAGCAGCAAGAGCAGCAAGAGCTTACCATAT	600
Db	185	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	204
QY	601	AGCCCATACAGCCCGCAAGCTAGCCTAGACAAGAAGAGCGTCAATTTAGCCCTCGAGGA	660
Db	205	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	224
QY	661	CAGCAGCGCGCAGAGAACGAGCAGCAAGAAGAGAAACCAAGGTGGAACATCTTC	720
Db	225	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe	244
QY	721	AGCGGCTTCACCGCGAGTTCCTGGAAACAGCCCTCCAGGTTACACACAGACAGATAGT	780
Db	245	SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal	264
QY	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACGTAGGGGA	840
Db	265	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	284
QY	841	GGCCTCAGAAATCTTGAGCCAGATAGAAGAGACGTGCCGACGAAGAAGAGATACGAT	900
Db	285	GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	304
QY	901	GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGG	960
Db	305	GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	324
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTCAAAAAGAACATTTGGTAGAAC	1020
Db	325	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	344
QY	1021	AGATCCCTGACATCTCAACACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC	1080
Db	345	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	364

QY	1081	CTTCTAATACTTTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA	1140
Db	365	LeuLeuIleLeuArgTyrLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	384
QY	1141	TTGTTTGTGCTCCTACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG	1200
Db	385	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	404
QY	1201	GCTCAGTCCAAGTGTGGACAGCAACGGCAACAGAGTGTACGACGAGAGCTTCAGAG	1260
Db	405	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	424
QY	1261	GGTCAAGTCTGTTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAGCGAGAAC	1320
Db	425	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	444
QY	1321	TTTCGAATACGTGGCATTTCAAGACAGACTCAAGCCCGCAGCATAGCCAACTCCCGGTGAA	1380
Db	445	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	464
QY	1381	AACTCGCTCATAGATAAAGTGGCGGAGGAGGTGGTGCAAATTCATATGGCTCCCAAGG	1440
Db	465	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	484
QY	1441	GAGCAGGCAAGGAGCTTAAGAACCAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG	1500
Db	485	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln	504
QY	1501	CAGTCTCCGAGGCTGTGGCT 1521	
Db	505	GlnSerProArgAlaValAla 511	
RESULT 10			
ADO38357			
ID	ADO38357	standard; protein; 507 AA.	
XX	ADO38357;		
XX	15-JUL-2004	(first entry)	
DT			
DE		Peanut allergen Ara h 3.	
XX		Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;	
KW		Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;	
KW		Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;	
KW		Antiallergic; Cytostatic; Antipsoriatic; Gene therapy; Vaccine;	
KW		MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;	
KW		influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;	
KW		rickettsia; rheumatoid arthritis; multiple sclerosis;	
KW		lupus erythematosus; diabetes mellitus; myasthenia gravis;	
KW		autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;	
KW		allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;	
KW		adenoma; peanut; Ara h 3.	
XX		Arachis hypogaea.	
OS			
XX		US2004058881-A1.	
PN			
XX		25-MAR-2004.	
PD			
XX		24-SEP-2002; 2002US-00253286.	
PF			
XX		24-SEP-2002; 2002US-00253286.	
PR			
XX		(ANTI-) ANTIGEN EXPRESS INC.	
PA			
XX		Humphreys RE, Xu M;	
PI			
XX		WPI; 2004-294259/27.	
DR			
XX			
PT		New non-naturally occurring protein or polypeptide modified by recombinant DNA techniques, useful for treating multiple sclerosis,	

PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
PT colitis, cancer or psoriasis.
XX

Example 3; Page 21-22; 90pp; English.

The invention relates to a non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques comprising: a C-terminal element comprising an MHC Class II-presented epitope; an N-terminal element comprising an Ii-key motif; and an intervening element comprising a sequence of 4-11 amino acid residues where the modification by recombinant DNA techniques taking place within elements (b) and (c). Also described are methods for: suppressing or enhancing an immune response directed toward an MHC (major histocompatibility complex) Class II-presented epitope of interest. Suppressing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding an Ii-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the Ii-key motif to decrease its conformance to the archetypal Ii-key regulatory motif. Enhancing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the nucleic acid sequence to introduce an Ii-key motif appropriately spaced from the MHC Class II-presented epitope. The protein or polypeptide of interest corresponds to a protein or polypeptide encoded by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza, preferably vaccinia virus. The non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques is useful for treating infectious diseases caused or associated with infection by a bacterium, virus, parasite, fungus, rickettsia or other infectious agents. It is also useful for treating rheumatoid arthritis, multiple sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis, autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma, allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or adenomas. The present sequence represents the amino acid sequence of peanut allergen Ara h 3 used in the invention.

SQ Sequence 507 AA;

Alignment Scores:
Pred. No.: 8e-231 Length: 507
Score: 2615.00 Matches: 498
Percent Similarity: 98.22% Conservative: 0
Best Local Similarity: 98.22% Mismatches: 9
Query Match: 97.18% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-3 (1-1524) x AD038357 (1-507)

Qy	1	CGGCAGCAACCGGAGGAGACGGCTGCGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
Db	1	ArgGlnProGluGluAsnAlaCysGlnPheGlyArgLeuAsnAlaGlnArgProAsp	20
Qy	61	AATGCGATTGAATCAGAGCGGGTTACATTAGACTTTGGAACCCCAACACGAGGATTC	120
Db	21	AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlyGluPhe	40
Qy	121	GAATGCGCGCGCTGCGCTCTCTGCTTAGTCTCCGCGCAACCGCCTTCGTAGCGCT	180
Db	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	60
Qy	181	TTCTACTTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATATTGGGTG	240
Db	61	PheTySerAsnAlaProGlnGluIlePheIleGlyGlyArgGlyTyrllePheGlyLeu	80
Qy	241	ATATTCCTGTTGTTCTAGACACTATGAAGAGCTTCACACAAAGTCTGCGATCTCAG	300
Db	81	IlePheProGlyCysProArgHisTyrlleGluProHisThrGlyGlyArgSerGln	100
Qy	301	TCCCAAGACCACCAAGACGTCTCCAGGAGAGACCAAGCCACAGCAACGAGATAGT	360

Db	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer	120
Qy	361	CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTTGCT	420
Db	121	HisGlnLysValHisArgPheAspGluGlyPheLeuIleAlaValProThrGlyValAla	140
Qy	421	TTCTGGCTCTACCAACGACGACGACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC	480
Db	141	PheTrpLeuTyrlleAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
Qy	481	AACAACGACAAACAGCTTGTATTCCTCCAGAGATTCATTTGGTGGGAAACACGAG	540
Db	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu	180
Qy	541	CAAGAGTCTTAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600
Db	181	GlnGluPheLeuArgTyrlleGlnGlnSerArgGlnSerArgArgSerLeuProTyrlle	200
Qy	601	AGCCCATACAGCCCGGCAAGTACCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	201	SerProTyrlleSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	220
Qy	661	CAGCAGACCGCAGACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
Db	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	240
Qy	721	AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTTCAGGTTGACGACAGACAGATAGT	780
Db	241	SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	260
Qy	781	CAAAACCTTAAGAGCGCAGACCGGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG	840
Db	261	GlnAsnLeuArgGlyGluThrSerGluSerGluGluGlyAlaIleValThrValArgGly	280
Qy	841	GGCTCAGAACTTTGAGCCAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	900
Db	281	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrllePhe	300
Qy	901	GAAGATGAATATGATATCGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	960
Db	301	GluAspGluTyrlleGluTyrlleAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
Qy	961	GGAAATGTTATGAAGAGAGAGATCTGCGACCGCAAGTCTTAAAGAAACATTTGTTAGAAC	1020
Db	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAsnIleGlyArgAsn	340
Qy	1021	AGATCCCTGACATCTACAAACCTCAAGCTGTTCTCACTCAAACTGCCAAGATCTCAAC	1080
Db	341	ArgSerProAspIleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
Qy	1081	CTTCTAATATCTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA	1140
Db	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArgAsnAla	380
Qy	1141	TTGTTTGTGCTCCTACTACCAACCAACGACGACGACGACGACGACGACGACGACGACG	1200
Db	381	LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArgGlyArg	400
Qy	1201	GCTCAGTGCAGCTGCTGGACACGACGACGACGACGACGACGACGACGACGACGACGACG	1260
Db	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluLeuGlnGlu	420
Qy	1261	GGTTCAGTGTGTTGTTGTCACAGAACTTTCGCGCTGCTGGAAAGTCCCAAGAGCGAGAAC	1320
Db	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
Qy	1321	TTGGAATAGTGGCATTCAGACAGACTCAAGGCCGACGACGACGACGACGACGACGACGAC	1380
Db	441	PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
Qy	1381	AACTCCGTTCATAGATAACCTCGCGAGGAGGTGTTGCAATTCATATTCGCTCCAAAGG	1440

```

Db      461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY      1441 GAGCAGCAAGGAGGCTTAAAGAACAAACACCCCTTCAAGTTCCTCGTTCCACCGTCTCAG 1500
Db      481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 500
QY      1501 CACTCTCCGAGGCGTGTGGCT 1521
Db      501 GlnSerProArgAlaValAla 507

RESULT 11
ABG71266
ID      ABG71266 standard; protein; 481 AA.
XX      AC      ABG71266;
XX      DT      17-DEC-2002 (first entry)
XX      DE      Glycine max (Soybean) var. Dare protein.
XX      KW      Soybean; Glycinin; atomic coordinate data; processability; soya protein;
KW      Dare; protein co-ordinate data.
XX      OS      Glycine max.
XX      JP2002193996-A.
XX      PN      10-JUL-2002.
XX      PD      21-DEC-2000; 2000JP-00405097.
XX      PR      21-DEC-2000; 2000JP-00405097.
XX      PA      (KYOU ) UNIV KYOTO.
XX      WP1; 2002-685438/74.
DR      N-PSDB; ABS55193.
XX      Glycinin, beta-conglycinin and proglycinin, their crystal structures,
PT      three dimensional coordinates, three dimensional structured and models
PT      and their uses.
XX      Disclosure; Page 1273-1274; 1298pp; Japanese.
XX      The present invention relates to a new Glycinin characterised by the
CC      atomic coordinate data fully defined in the specification. The structure
CC      can be used for improving processability of soya protein. The present
CC      amino acid sequence represents the Glycine max (Soybean) var. Dare
CC      protein, as described in the specification
XX      SQ      Sequence 481 AA;

Alignment Scores:
Pred. No.:      1.32e-128      Length:      481
Score:          1502.00      Matches:      297
Percent Similarity: 71.57%      Conservative: 68
Best Local Similarity: 58.24%      Mismatches: 91
Query Match:      55.82%      Indels:      54
DB:              5      Gaps:      7

US-10-728-323-3 (1-1524) x ABG71266 (1-481)
QY      1 CGGCAGCAACGGAGGAGAACGGTCCAGTTCAGCGCTCAATGCGCAGAGACCTGAC 60
Db      23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY      61 AATCGCATGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACCAACGAGGATTC 120
Db      43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY      121 GAATGCGCGCGCTCGCTCTCTCGCTTAGTCTCTCGCGCGCAACCGCTTCGTAGGCT 180
Db      63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82

```

```

QY      181 TTCTACTCAATGCTCCCGCAGGAGATCTTTCATCCAGCAAGGAGGATACCTTTGGGTTG 240
Db      83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY      241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCACACACAGGTCGTGATCTCAG 300
Db      103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY      301 TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGAACCAAGCCACACAGCAACGAGATAGT 360
Db      120 SerSerArgPro-----GlnAspArg 126
QY      361 CACCAGAGGTGCACCGCTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGTGT 420
Db      127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY      421 TTCTGGCTCTACAAGCACCAGCAGACTGATGTTGCTGTTCTCTTACTGACACCAAC 480
Db      147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY      481 AACAAACGACACACAGCTTGATCAGTTCCTCCAGGAGATTCAATTGGCTGGGACACGGAG 540
Db      167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY      541 CAAGAGTCTTAAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db      187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY      601 AGCCCATACAGCCCGCAAGTCAAGCTAGACAGAAAGAGCGTGAATTTAGCCCTCGAGGA 660
Db      200 -----Gly 200
QY      661 CAGCAGACCGCAGAGAACGAGCAGACAGAAAGAAAGAAAGAGTGTGAACATCTTC 720
Db      201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlySerIleLeu 220
QY      721 AGCGGCTTCACCGCGAGTTCCTGGAACAAAGCTTCCAGGTTCGACGACAGACAGATAGTG 780
Db      221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY      781 CAAAACCTTAAGAGGCGAGACCGCAGAGTGAAGAGAGGAGGCCATTGTGACAGTACGGGA 840
Db      240 ArgLysLeuGlnGlyLysGlnGluGluLysGlyValAlaIleValThrValLysGly 259
QY      841 GCGCTCAGAACTCTGAGCCCA-----GATAGAAAGAGAGCTGCCACCAAGAGAG 891
Db      260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
QY      892 GAATACGATGAAGATGAATATATGATACGATGAAGAGGATGAAGGCGTGCAGGGGAAGC 951
Db      280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY      952 AGAGCAGGGGGAATGTTATTAAGAGACGATCTCGACCGCAAGTGTCTAAAAGAACAT 1011
Db      295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY      1012 GGTAGAAACAGATCCCTGACATCTACACCTCAGCTGAGTGTCTCACTAAAAGTGCACAC 1071
Db      312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
QY      1072 GATCTCAACCTCTAATCTAGTGGCTTGGACCTTAGTGTGATATATGGAATCTCTAC 1131
Db      332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY      1132 AGAATGATGTTGTTGCTGCTCACTACACCAACGACGACAGCATCATATATCATGTTG 1191
Db      352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY      1192 AGGGGACGGCTCAGCTGCAAGTGTGGACAGACACGCAACGAGGTACACAGGAGGAG 1251
Db      372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391

```


Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
Qy 121 GAATCGCGCGCTCGCCCTCTCTCGTTAGTCTCCGCGCAACAGCGCCCTCTGTAGCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
Qy 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATATCTTTGGGTG 240
Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 102
Qy 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCTCACACAAAGGTCTGCTATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
Qy 301 TCCCAAGACACCAAGAGCTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
Qy 361 CACCAGAGGTGACCGCTTTCATGAGGTGATCTCATTCAGTTCCACCGGTGTGCT 420
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
Qy 421 TTCTGGCTCTACACGACCGACACTGATGTGTCTGCTGTTCTCTTACTGACCAAC 480
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
Qy 481 AACACACACACAGCTTGATCAGTCCCGAGGAGATCAATTGGTGGGAAACAGGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
Qy 541 CAAGAGTCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAGAAGCTTACCATAT 600
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 199
Qy 601 AGCCCATACAGCCCGCAAGTACGCTTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 200 -----Gly 200
Qy 661 CAGCAGACGCGCAGAGAACGAGCAGGAGCAAGAGAAACGAAAGTGAACATCTTC 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 220
Qy 721 AGCGCTTACCGCGGAGTCTCGACACAGCTTCCAGGTTGACGACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
Qy 781 CAAACCTTAAGAGCAGACCGAGAGTGAAGAGAGCGCATTTGTGACAGTGAGGGGA 840
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly 259
Qy 841 GGCTCAGATCTTGACCCA-----GATAGAAAGAGAGCTGCCGACGAAGAGAG 891
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
Qy 892 GAATACGATCAAGATGATATGATACGATCAAGAGATAGAGCGGTGGCAGGGAGAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
Qy 952 AGAGCGAGGGGGAATGTTATTAAGAGACGATCTGCACCCCAAGTGTCTAAAGAAACATT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
Qy 1012 GGTAGAACAGATCCCTGACATCTACACCTCAAGCTGGTTCATCTCAAAATGCCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
Qy 1072 GATCTCAACCTTCTAATCTAGTGGCTGGACCTAGTCTGAAATGGAATCTCTAC 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
Qy 1132 AGGAATGTCATTGTTGTGCTACTACACCAACCAAGCAGACATCATATATCATG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371

Qy 1192 AGGGGACGGGCTCACGTGCAAGTCTGGAGACGCAACAGAGTGTACGACGAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
Qy 1252 CTTCAAGAGGGTCCAGCTGCTTGTGGTCCACAGAACTTCGCCCTCGCTGGAAAGTCCCAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
Qy 1312 AGCGAGAATCTCGAATACGTGGCATTCACAGACAGACTCAAGGCCAGCATAGCAACCTC 1371
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
Qy 1372 GCCGTGAAAACTCCGTCTCATAGATAACCTGCCGAGAGAGGTGTTGCAAAATTCATATGGC 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
Qy 1432 CTCCTCAAGGGAGCAGGCAAGGAGCTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCA 1491
Db 452 LeuArgArgGlnAlaAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
Qy 1492 CCCTCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db 472 ProLysGluSerGlnArgArgValValAla 481

RESULT 14

ADG43988
ID ADG43988 standard; protein; 481 AA.

AC ADG43988;

DT 26-FEB-2004 (first entry)

DE G. max glycinin G3 subunit protein.

KW oil content; plant; storage protein; seed-specific promoter; 2S-albumin;
KW 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;
KW oil production; fat production; free fatty acid production; food;
KW animal feed; pharmaceutical; fine chemical production; glycinin.

OS Glycine max.

PN WO2003077643-A2.

PD 25-SEP-2003.

XX 17-MAR-2003; 2003WO-BP002733.

XX 20-MAR-2002; 2002DE-01012893.

XX (BADI) BASF PLANT SCI GMBH.

XX Bauer J;

XX WPI; 2004-011485/01.

XX N-PSDB; ADG43987.

PT Increasing total oil content of plants, useful e.g. as foods or animal
PT feeds, by reducing amount of storage proteins, particularly with double-
PT stranded interfering RNA.

PS Claim 4; SEQ ID NO 28; 253pp; German.

XX This invention describes a novel method for increasing the total oil
CC content of a plant by reducing the amount of at least one storage protein
CC in the plant (or its tissue, organs, parts or cells) and selecting plants
CC that have higher total oil content than starting plants. The storage
CC protein is suppressed by introducing antisense RNA, optionally combined
CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding
CC factors directed against storage protein genes, viral sequences that
CC degrade storage protein RNA, constructs that induce homologous
CC recombination of endogenous storage protein genes or mutations into
CC storage protein genes. Most preferably a plant cell is stably transfected

CC with a recombinant expression construct, then regenerated to plants that
 CC express the incorporated sequence. The expression constructs particularly
 CC contain a seed-specific promoter and they are introduced into plants by
 CC standard methods, e.g. via Agrobacterium. The preferred storage proteins
 CC of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-
 CC prolamines. Transgenic organisms produced by the new method are used for
 CC production of oils, fats, free fatty acids or their derivatives, useful
 CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence
 CC represents a storage protein used to illustrate the method of the
 CC invention.

XX SQ Sequence 481 AA;

Alignment Scores:

Pred. No.: 1 32e-128 Length: 481
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.57% Conservatives: 68
 Best Local Similarity: 58.24% Mismatches: 91
 Query Match: 55.82% Indels: 54
 DB: 8 Gaps: 7

US-10-728-323-3 (1-1524) x ADG43988 (1-481)

QY 1 CGCAGCAACCGGAGGAGACCGGCTCCAGTTCAGCGCTCAATCGCAGACCTGAC 60
 Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
 QY 61 AATCGCATTAATCAGAGGCGGTACATTGACACTTGGNACCCCAACACGAGGATTC 120
 Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
 QY 121 GAATGGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
 Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
 QY 181 TTCTACTCCATGCTCCCCAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTG 240
 Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
 QY 241 ATATTCCTGCTGCTCCTAGACTATGAGAGCTTCACACAAAGTCTGCTCATCTCAG 300
 Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
 QY 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
 Db 120 SerSerArgPro-----GlnAspArg 126
 QY 361 CACCAAGAGTGACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTGCT 420
 Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
 QY 421 TTCTGGCTCTACACGACGACGACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
 Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValAlaValSerLeuIleAspThrAsn 166
 QY 481 AACACGACCAACCGCTTGATCTAGTTCCCGCAGGAGATTCAAATTTGGTGGGAACGGAG 540
 Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
 QY 541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATAT 600
 Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
 QY 601 AGCCCATACGCGGCAAGTACGCTCAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
 Db 200 -----Gly 200
 QY 661 CAGCAGACGCGCAGAGAACGAGCAGGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 QY 721 AGCGGCTTACGCGGAGTTCTTCGGAACAAGCCCTTCAGGTTGACGACGACAGATAGTG 780

Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
 QY 781 CAAACCTAAGAGGCGGAG 840
 Db 240 ArgLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 259
 QY 841 GGCCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCGTCCGACGAAAGAGAG 891
 Db 260 GlyLeuSerValIleSerProThrGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlu 279
 QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGGATGAAGAGCGTGGCAGGGAAGC 951
 Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
 QY 952 AGAGCGAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTGCTAAAAGACATT 1011
 Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
 QY 1012 GGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGTTCACATAAACTGCCAAC 1071
 Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
 QY 1072 GATCTCAACCTTCTAATCTAGTGGCTTGGACCTAGTCTGCTGAATATGAAATCTCTAC 1131
 Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
 QY 1132 AGAATGCACTTGTTCCTCCTACACACCAACGACGACGACGATCATATATCGATTG 1191
 Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
 QY 1192 AGGGGACGGGCTCAGTGCAGCTCGTGGAGCAGCAACGCGCAACGAGTGTACGACGAGAG 1251
 Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
 QY 1252 CTTCAAGAGGTGACGTGCTTGTGTCGACAGAACTTCGCGTCTGCTGAAAGTCCAG 1311
 Db 392 LeuGlnGluGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
 QY 1312 ACGGAACTTCGAATACGCTTCAAGACGACTCAAGCCGCGGAGGAGTGTGCAATTCATATGC 1371
 Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
 QY 1372 GCCGTGAAAACTCCGTCATAGATAACCTCCGCGGAGGAGTGTGCAATTCATATGC 1431
 Db 432 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
 QY 1432 CTTCAAGGAGGACGAGCAGGAGCTTAAGAACACAAACCCCTTCAAGTCTCTCGTTCCA 1491
 Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
 QY 1492 CGGTCTCAGCAGTCTCCGAGGGCTGGCT 1521
 Db 472 ProLysGluSerGlnArgArgValValAla 481
 RESULT 15
 ABG71265
 ID ABG71265 standard; protein; 485 AA.
 XX AC ABG71265;
 DT 17-DEC-2002 (first entry)
 XX DE Glycine max (Soybean) var. Shiotsurunoko protein #2.
 XX KW Soybean; Glycinin; atomic coordinate data; processability; soya protein;
 XX KW Shiotsurunoko; protein co-ordinate data.
 XX OS Glycine max.
 XX PN JP2002193996-A.
 XX PD 10-JUL-2002.
 XX

PF 21-DEC-2000; 2000JP-00405097.
 XX
 PR 21-DEC-2000; 2000JP-00405097.
 XX
 PA (KYOU) UNIV KYOTO.
 XX
 DR WPI; 2002-685438/74.
 DR N-PSDB; ABS55192.
 XX
 PT Glycinin, beta-conglycinin and proglycinin, their crystal structures,
 PT three dimensional coordinates, three dimensional structured and models
 PT and their uses.
 XX
 PS Disclosure; Page 1269-1271; 1298pp; Japanese.
 XX
 CC The present invention relates to a new Glycinin characterised by the
 CC atomic coordinate data fully defined in the specification. The structure
 CC can be used for improving processability of soya protein. The present
 CC amino acid sequence represents the Glycine max (Soybean) var.
 CC Shirotsubunoko protein #2, as described in the specification
 XX
 SQ Sequence 485 AA;

Alignment Scores:
 Pred. No.: 2,25e-128 Length: 485
 Score: 1499.50 Matches: 297
 Percent Similarity: 72.02% Conservative: 71
 Best Local Similarity: 58.12% Mismatches: 94
 Query Match: 55.72% Indels: 49
 DB: 5 Gaps: 6

US-10-728-323-3 (1-1524) x ABG71265 (1-485)

QY 1 CGGAGCAACGGAGAGACGCGTCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60
 DB
 20 ArgGluGlnAlaGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 39
 QY 61 AATCGCATTCAGTACAGAGGGCGGTACATTGAGACTTGGAACTCCCAACACACAGGAGTTC 120
 DB
 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 59
 QY 121 GAATCGCGCGCGCGCTCTCTCTGCTTAGTCTCCGCGCAGACGCGCTTCGTAGGCGCT 180
 DB
 60 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 79
 QY 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGGTTG 240
 DB
 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
 QY 241 ATATTCCCTGGTGTCTAGACACTATGAGAGGCTCAGACACAGAGTCTGATCTCAG 300
 DB
 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 119
 QY 301 TCCCAAGACACCAAGACGCTCTCCAGGAGAGACCAAGCAACAGCAACGAGATAGT 360
 DB
 120 SerGlnArgPro-----GlnAspArg 126
 QY 361 CACCAGAGGTGACCGGTTTCGATGAGGGTGATCTCATTCAGTCTCCACCGGTGTGCT 420
 DB
 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
 QY 421 TTCTGGCTCTACACGACACGACGACTGATGTTGTGCTGTTCTCTTCTTACTGACCAAC 480
 DB
 147 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 166
 QY 481 AACCAACACACACGCTTATCATCTCCAGGAGATTCAATTGGCTGGGACACGGAG 540
 DB
 167 SerLeuGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
 QY 541 CAAGAGTCTTAAAGTACACGACCAACAAAGCAGACCAAGCAGACCAAGAGCTTACCATAT 600
 DB
 187 GlnGluPheLeuLysTyrGlnGlnGln-----

Search completed: August 24, 2005, 09:56:21
 Job time : 110.311 secs

QY 601 AGCCCATACAGCCCGCAAGTCTAGACAAAGAGCGGTGAATTTAGCCCTTCGAGGA 660
 DB
 196 -----GlnGlnGlyGlySerGlnSerGlnLysGly 205
 QY 661 CAGCACACCGCAGAGAACGAGCAGACAGAAAGAAAGAAAGAAAGTGAACATCTTC 720
 DB
 206 LysGln-----GlnGluGluAsnGluGlySerAsnIleLeu 218
 QY 721 AGCGGCTTACGCGGAGTCTTGGAAACAGCCCTTCAGGTTACACAGACAGATAGTG 780
 DB
 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
 QY 781 CAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGCGCCATTGTGACAGTGGGGA 840
 DB
 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
 QY 841 GGCTCAGAAATCTTGAGCCAGATAGAAAGAGACGTCGCCACCAAGAAAGAGATACGAT 900
 DB
 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 277
 QY 901 GAAGATGAA-----TATGAATACGATGAAGAGAGATAGAGCGGTGCGAGGGA 948
 DB
 278 GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
 QY 949 AGCAGAGCGAGGGGAAATGATTGAAGAGACGATCTGCACCAAGTGTATAAAGAAC 1008
 DB
 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
 QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGTTTCACTCAAACTGCC 1068
 DB
 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
 QY 1069 AAGCATCTCAACCTTCTAATACCTAGTGGCTTGGACCTAGTCTGAATATGAAATCTC 1128
 DB
 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
 QY 1129 TACAGAAATGCTTGTCTGCTCACTACACACCAAGCAGCAGCATCATATATCGA 1188
 DB
 355 ArgLysAsnAlaMetValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
 QY 1189 TTGAGGGGACGGGCTCAGCTGCAAGTCTGAGCAGCAACGGCAACAGAGTGTACGACGAG 1248
 DB
 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394
 QY 1249 GAGCTTCAAGAGGTCAGCTGTGTGTGTCAGCAAGAACTTCGCTCGCTGAAAGTCC 1308
 DB
 395 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 414
 QY 1309 CAGAGCGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAAC 1368
 DB
 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
 QY 1369 CTCGCGGTGAAACTCTCGTCATAGATAACCTGCGGAGAGGTGTCAAATTCATAT 1428
 DB
 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
 QY 1429 GGCCTTCAAAGGAGCAGCAAGCGCAGTCTAAACAACAACACCCCTTCAAGTTCTTCGT 1488
 DB
 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
 QY 1489 CCACGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
 DB
 475 ProProGlnGluSerGlnArgArgAlaValAla 485

Result No.	Score	Query Match	Length	DB ID	Description
1	1502	55.8	481	2	S04605
2	1499.5	55.7	485	2	S11002
3	1496	55.6	495	2	S10851
4	1492.5	55.5	485	1	FW5VG1
5	1466	54.5	495	1	FW5VG2
6	1457.5	54.2	498	2	S44294
7	1455	54.1	520	2	S08237
8	1440	53.5	517	1	FWPMLA
9	1431.5	53.2	482	2	S48677
10	1428	53.1	484	2	S11003
11	1414.5	52.6	500	2	S14393
12	1384	51.4	507	2	T08452
13	1375	51.1	497	2	S14392
14	1073.5	39.9	551	2	S51941
					glycinin G3 - soyb
					glycinin G2 precu
					glycinin G1 precu
					glycinin chain A2B
					glycinin chain A1a
					legumin A precursor
					legumin A2 precursor
					legumin A precursor
					legumin A precursor
					glycinin G3 precu
					legumin A2 precursor
					probable legumin A
					legumin A1 precursor
					prunin 1 precursor

```
121 GAATGCGCGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
   ::::::::::::::::::::::::::::
63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
   ::::::::::::::::::::::::::::::
181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCAGCAAGGAGGATCTTTGGTTG 240
   ::::::::::::::::::::::::::::::
83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
   ::::::::::::::::::::::::::::::
241 ATATTCCCTCGTTGTCTAGACACTATGAAGAGCCTCACACACAAAGTCTCGATCTCAG 300
   ::::::::::::::::::::::::::::::
103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
   ::::::::::::::::::::::::::::::
301 TCCCAAGACACCACAAAGACGCTCTCCAAGGAGAAACCAAGCCACAGCAACGAGATAGT 360
   ::::::::::::::::::::::::::::
120 SerSerArgPro-----GlnAspArg 126
   ::::::::::::::::::::::::::::
361 CACAGAGAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTTCCACCGGTGTGCT 420
   ::::::::::::::::::::::::::::
127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
   ::::::::::::::::::::::::::::
421 TTCTGGCTCTACACGACACGACACTGATGTGTGTGCTGTTCTTCTTACTGACACCAAC 480
   ::::::::::::::::::::::::::::
147 TyrIrrMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
   ::::::::::::::::::::::::::::
481 AACACGACACACGAGCTTGATAGTTTCCCGAGGAGATTCAATTGGCTGGGAACACGGAG 540
   ::::::::::::::::::::::::::::
167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
   ::::::::::::::::::::::::::::
541 CAAGAGTTCTTAGGTACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATAT 600
   ::::::::::::::::::::::::::::
187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
   ::::::::::::::::::::::::::::
601 AGCCCATACAGCCCGCAAGTCAAGCTAGACAGAGAGCGTGAATTAGCCCTCGAGGA 660
   ::::::::::::::::::::::::::::
200 -----Gly 200
   ::::::::::::::::::::::::::::
661 CAGCACAGCCGACAGACGAGCAGGACGAGCAAGAAAGAAACGAAGTGGAAACATCTTC 720
   ::::::::::::::::::::::::::::
201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 220
   ::::::::::::::::::::::::::::
721 AGCGGTTTCAGCGCGGAGTTCTCGAACAGCCCTTCAGGTTTACACAGACAGATAGTG 780
   ::::::::::::::::::::::::::::
221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
   ::::::::::::::::::::::::::::
781 CAAACCTTAAGCGCAGACCGGAGAGTGAAAGAGGAGCCATTGTGACAGTGAGCGGA 840
   ::::::::::::::::::::::::::::
240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly 259
   ::::::::::::::::::::::::::::
841 GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGACGTCGCCGACGAAGAAGAG 891
   ::::::::::::::::::::::::::::
260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
   ::::::::::::::::::::::::::::
892 GAATACGATCAAGATGAATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 951
   ::::::::::::::::::::::::::::
280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
   ::::::::::::::::::::::::::::
952 AGAGCGCGGGGAATGTTATTGAAGAGACGATCTGCACGCAAGTCGTAATAAGAACATT 1011
   ::::::::::::::::::::::::::::
295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
   ::::::::::::::::::::::::::::
1012 GGTAGAAACAGATCCCTGATCACTCAACCCCTCAAGCTGGTTTCACTCAAAACTGCCAAC 1071
   ::::::::::::::::::::::::::::
312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
   ::::::::::::::::::::::::::::
1072 GATCTCAACCTTCTAATAGTGTGGCTGGACCTAGTGTGCTGAATATGGAATCTCTAC 1131
   ::::::::::::::::::::::::::::
332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
   ::::::::::::::::::::::::::::
1132 AGGAATGCATTTGTTGCGCTCACTACAACCAACGACACACAGCATCATATATCATG 1191
   ::::::::::::::::::::::::::::
352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
   ::::::::::::::::::::::::::::
1192 AGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGCGCAACAGAGGTGTACGAGGAG 1251
   ::::::::::::::::::::::::::::
```

```
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
   ::::::::::::::::::::::::::::
QY 1252 CTTCAAGAGGTCAGCTGCTTCTGTCGTCACAGAACTTCGCGCTCGCTCGAAGTCCCAAG 1311
   ::::::::::::::::::::::::::::
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
   ::::::::::::::::::::::::::::
QY 1312 ACCGAGAACTTCGAATACCTGTCATTAAGACAGACTCAAGGCCCAAGCATAGCCACCTC 1371
   ::::::::::::::::::::::::::::
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
   ::::::::::::::::::::::::::::
QY 1372 GCGGTGAAAACCTCCGTCATAGATAACCTGCGGAGGAGGTGGTTCGCAAAATTCATATGC 1431
   ::::::::::::::::::::::::::::
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
   ::::::::::::::::::::::::::::
QY 1432 CTTCAAGGAGCAGCGCAGGAGCTTAAAGACACACACCCCTTCAAGTTCCTGTTCCA 1491
   ::::::::::::::::::::::::::::
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
   ::::::::::::::::::::::::::::
QY 1492 CCGTCTCAGCAGCTCCGAGGCGCTGGCT 1521
   ::::::::::::::::::::::::::::
Db 472 ProLysGluSerGlnArgArgValValAla 481
   ::::::::::::::::::::::::::::

RESULT 2
S11002
Glycinin G2 precursor - soybean
N:Alternate names: glycinin A2B1a
C:Species: Glycine max (soybean)
C:Date: 21-Nov-1993 #sequence revision 19-Jan-1996 #text_change 09-Jul-2004
C:Accession: S11002; S04604; A26990
R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;
Plant Cell 1, 313-328, 1989
A:Title: Characterization of the glycinin gene family in soybean.
A:Reference number: S10851; MUID:92393391; PMID:2485233
A:Accession: S11002
A:Molecule type: DNA
A:Residues: 1-485 <NIE>
A:Cross-references: UNIPROT:P04405
A:Experimental source: variety Dare
R:Thanh, V.H.; Tumer, N.E.; Nielsen, N.C.
Nucleic Acids Res. 17, 4387, 1989
A:Title: The glycinin Gy(2) gene from soybean.
A:Reference number: S04604; MUID:89296499; PMID:2740230
A:Accession: S04604
A:Molecule type: DNA
A:Residues: 1-485 <THA>
A:Cross-references: EMBL:X15122; NID:g18636; PIDN:CRA33216.1; PID:g18637
R:Fukazawa, C.; Momma, T.; Higuchi, W.; Uda, K.
Nucleic Acids Res. 15, 8117, 1987
A:Title: Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit prec
A:Reference number: A26990; MUID:88040439; PMID:3671077
A:Accession: A26990
A:Molecule type: DNA
A:Residues: 1-485 <FUK>
A:Cross-references: GB:X02806
C:Genetics:
A:Gene: Gy2
A:Introns: 93/1; 177/3; 356/3
C:Superfamily: glycinin
C:Keywords: storage protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-485/Product: glycinin G2 #status predicted <MAT>

Alignment Scores:
Pred. No.: 1,578-103 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.02% Conservative: 71
Best Local Similarity: 58.12% Mismatches: 94
Query Match: 55.72% Indels: 49
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x S11002 (1-485)
```

Qy	1	CGGCAGCAACCGGAGGAGAAACCGCTGCGAGTTCAGCGCTTCAATGCGCAGAGACCTGAC	60
Db	20	ArgGluGlnAlaGlnAsnGluCysGlnIleGlnLysLeuAenAlaLeuLysProAsp	39
Qy	61	AATCGCATTCGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGGATTC	120
Db	40	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAenAnLysProPhe	59
Qy	121	GAATGCGCCGCGCTGCGCCTCTCTCGGTTAGTCTCCGCGCAACGCGCTTCGTAGGCGCT	180
Db	60	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	79
Qy	181	TTCTACTCCAATGCTCCCGCAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTG	240
Db	80	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet	99
Qy	241	ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTCGTCTCGATCTCAG	300
Db	100	IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnAlaArgGlyArg	119
Qy	301	TCCCAAGACACCAAGACGCTCCAAGGAGAGAACCAACCAAGCAACAGCAGATAGT	360
Db	120	SerGlnArgPro-----GlnAspArg	126
Qy	361	CACCAGAAGTGCAACGTTTCGATGAGGTCATCTCATTCAGTTCGCCCGGTGTGCT	420
Db	127	HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	146
Qy	421	TTCTGGCTCTACACGACACGACACTGATGTGTGCTGCTTCTCTTACTGACACCAAC	480
Db	147	TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn	166
Qy	481	AACAACACCAACCAAGCTTGATTCAGTCTCCAGGAGATTCAATTGGCTGGCAACACGGAG	540
Db	167	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
Qy	541	CAAGAGTTCCTTAAGGTACCACGACCAACAAGCAGACAAAAGCAGACAGCAAGAGCTTACCATAT	600
Db	187	GlnGluPheLeuLysTyrGlnGln-----GlnGlnGlySerGlnSerGlnLysGly	205
Qy	601	AGCCCATACAGCCGCAAGATCAGCTCCTAGACAAAGAGAGCGTGTAATTAGCCTCCGAGGA	660
Db	196	-----GlnGlnGlySerGlnSerGlnLysGly	205
Qy	661	CAGCACAGCCGACAGAAACGACGAGGACAAAGAAAGAAACGAAGTGGAAACATCTTC	720
Db	206	LysGln-----GlnGlnGluAsnGluGlySerAsnIleLeu	218
Qy	721	AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACACAGATAGTG	780
Db	219	SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal	237
Qy	781	CAAACTTAGAGCGGACGACGAGAGTGAAGAAGAGGGGACCATTTGACAGTGAAGGGGA	840
Db	238	ArgAsnLeuGlnGlyLysAsnGluGluGluAspSerGlyAlaIleValThrValLysGly	257
Qy	841	GGCCTCAGAATCTTGACCCAGATAGAAAGAGAGCTGCCGACGAAGAGAGGAATACGAT	900
Db	258	GlyLeuArgValThrAlaProAlaMetCysArgLysProGlnGlnGluAlaAspAspAsp	277
Qy	901	GAAGATGAA-----TATGAATACATGAAGAGGATGAAGCGCTGGCGGGGA	948
Db	278	GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnAlaArgGlnSerLysArg	297
Qy	949	AGCAGAGCGGCGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAAC	1008
Db	298	SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn	314
Qy	1009	ATTGGTAGAAACAGATCCCTGACATCTACACCTCAAGCTGGTTCACATCAAAACCTGCC	1068
Db	315	IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrThrAla	334

QY	1069	AACGATCTCAACCTCTTAATACTTAGTGGCTGGACCTAGTGCTCAAGTAATATGGAATCTC	1128
DB	335	ThrSerLeuAspPheProAlaLeuTyrPheLeuLeuSerAlaGlnTyrGlySerLeu	354
QY	1129	TACAGAGAATGCAATGTTGTCGCTCACTCAACACCAACGACACAGCATCATATATCGA	1188
DB	355	ArgGlyAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla	374
QY	1189	TTGAGGGGACGGGCTCACGTGCAAGTCGTGGACACCAACGCGCAACAGAGTGTACGACGAG	1248
DB	375	LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly	394
QY	1249	GAGCTTCAAGAGGGTCACGTGCTTGTGGTGGCCACAGAACTTCGCCGTCGCTGGAAGTCC	1308
DB	395	GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaAlaLysSer	414
QY	1309	CAGAGCGAGAACTTCGAATACGTGGCATTCAGACACAGACTCAAGGGCCAGCATAGCCAAC	1368
DB	415	GlnSerAspAsnPheGluTyrValSerPheLeuThrAsnAspArgProSerIleGlyAsn	434
QY	1369	CTCGCGGTGMAAATCTCCGTCTAGATAACCTCCGCGGAGGAGGTGGTTGCAAAATTCATAT	1428
DB	435	LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe	454
QY	1429	GGCTCTCAAGGAGGACGAGGACAGCTTAAAGAACAAACACCCCTTCAAGTCTTCGTT	1488
DB	455	AsnLeuLysSerGlnGlnAlaArgGlnValValAsnAsnAsnProPheSerPheLeuVal	474
QY	1489	CCACCTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
DB	475	ProProGlnGluSerGlnArgArgAlaValAla	485
RESULT 3			
S10851			
glycinin G1 precursor - soybean			
N;Alternate names: Glycinin AlaBx			
N;Contains: glycinin chain Ala; glycinin chain Bx			
C;Species: Glycine max (soybean)			
C;Date: 21-Nov-1993 #sequence revision 19-Jan-1996 #text_change 09-Jul-2004			
C;Accession: S10851; S04603; JS0015			
R;Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fischer, Plant Cell 1, 313-328, 1989			
A;Title: Characterization of the glycinin gene family in soybean.			
A;Reference number: S10851; MUID:92393391; PMID:2485233			
A;Accession: S10851			
A;Molecule type: DNA			
A;Residues: 1-495 <NIE>			
A;Cross-references: UNIPROT:P04776			
A;Experimental source: variety Dare			
R;Sims, T.L.; Goldberg, R.B.			
Nucleic Acids Res. 17, 4386, 1989			
A;Title: The glycinin Gy(1) gene from soybean.			
A;Reference number: S04603; MUID:89296498; PMID:2740229			
A;Accession: S04603			
A;Molecule type: DNA			
A;Residues: 1-495 <SIM>			
A;Cross-references: EMBL:X15121; NID:g18634; PIDN:CAA33215.1; PID:g18635			
A;Experimental source: variety Dare			
R;Utsuni, S.; Kohno, M.; Mori, T.; Kito, M.			
J. Agric. Food Chem. 35, 210-214, 1987			
A;Title: An alternate cDNA encoding glycinin Ala Bx subunit.			
A;Reference number: JS0015			
A;Accession: JS0015			
A;Molecule type: mRNA			
A;Residues: 1-495 <UTS>			
A;Experimental source: var. Shiotsurunoko			
C;Genetics:			
A;Gene: Gyl			
A;Introns: 96/1; 180/3; 366/3			
C;Superfamily: glycinin			
C;Keywords: storage protein			
F;1-19/Domain: signal sequence #status predicted <SIG>			
F;20-495/Product: glycinin G1 #status predicted <NATI>			

F;20-310/Product: glycinin Ala chain #status predicted <MAT2>
F;311-495/Product: glycinin Bx chain #status predicted <MAT3>

Alignment Scores:

Pred. No.: 2,87e-103 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.87% Conservative: 73
Best Local Similarity: 56.70% Mismatches: 100
Query Match: 55.59% Indels: 50
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x S10851 (1-495)

QY	1	CGGACGCAACCGGAGGACCGCTGCCAGTTCAGCGCCTCAATGCGAGACCTGAC	60
DB	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGCATGAATCAGAGCGCGTTACATTGAGACTTGGAAACCCCAACACCGAGATTTC	120
DB	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTriAsnProAsnAsnLysProPhe	62
QY	121	GAATGGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGCCT	180
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	82
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTTGGTTG	240
DB	83	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet	102
QY	241	ATATTCCTCGTTGTCTAGACACTATGAAGACCTCACACAAAGGTTCGTGATCTCAG	300
DB	103	IleTyrProGlyCysProSerThrPheGluGluProGlnGlnArgGlyGln	122
QY	301	TCCCAAGACCACCAAGACTCTCCAAGGAGAGACCAAGCCCAACAGCAAGAGATAGT	360
DB	123	SerSerArgPro-----GlnAspArg	129
QY	361	CACGAGAGGTGCACCGTTTCGATAGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT	420
DB	130	HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	149
QY	421	TTCTGCTCTACACGACCAACGACACTGATGTTGTCTGTTTCTCTACTGACACCAAC	480
DB	150	TriTriMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn	169
QY	481	AACACGACACACGCTTGATGATTCCTCCCGAGAGATTCATTTGGCTGGGAACGGAG	540
DB	170	SerLeuGluAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu	189
QY	541	CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGACTTTACCATAT	600
DB	190	GlnGluPheLeuLysTyrGlnGlnGluGln-----	199
QY	601	AGCCCATACAGCCCGAAAGTACGCTAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660
DB	200	-----GlyGly	201
QY	661	CACGACGCGCAGACNACGACGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB	202	HisGlnSerGlnLysGlyLysHisGlnGlnGluGluGluAsnGluGlyGlySerIleLeu	221
QY	721	AGCGGCTTCACGCGGAGTTCTCTGGAAACAGCCCTTCAGGTTTCAGCAGACAGACAGATGTG	780
DB	222	SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla	240
QY	781	CAAAACCTAAGAGCGGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840
DB	241	LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly	260
QY	841	GGCCTCAGATCTTGACCCA-----GATAGAAAGAGAGCTGCCACGAGAGAGAG	891
DB	261	GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu	280

QY	892	GAATACGATGAAGATGAATATCAATAC-----GATCAAGAG---GATAGAAGG	936
DB	281	GluGluGluGluAspGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro	300
QY	937	CCTGCGAGGGGAAGCAGAGCGGAGGGAATGTTATTGAAGACGATTCGACCGCAAGT	996
DB	301	ArgGlySerGlnSerLysSerArgAsnGlyIleAspGluThrIleCysThrMetArg	320
QY	997	GCTAAAGAACATTTGTTAGAAACAGATCCCTGCATCTACAAACCTCAACCTGGTTCA	1056
DB	321	LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer	340
QY	1057	CTCAAAATGCCCAACGATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTGTGAA	1116
DB	341	ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTriPLeuArgLeuSerAlaGlu	360
QY	1117	TATGGAATCTCTACAGGAATGCATTTGTTGCGTCTACTACAACACCAACGACACAGC	1176
DB	361	PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer	380
QY	1177	ATCATATATCGATTGAGGGGACGGCTCACGTGCGCAAGTCGTGCAGCAGCAACGCAACAG	1236
DB	381	IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValValAsnCysAsnGlyGluArg	400
QY	1237	GTGTACGACGAGGAGCTTCAGAGGGTCCAGGTGCTTGTGGTCCACAGAACTTCGCCGTC	1296
DB	401	ValPheAspGlyLeuLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValVal	420
QY	1297	GCTGGAAGTCCGAGCGGAGAACTTCGAATACGTGGCATTCACAGACAGACTCAAGGCC	1356
DB	421	AlaAlaArgSerGlnSerAspAsnGluTyrValSerPheLysThrAsnAspThrPro	440
QY	1357	AGCATAGCAACCTCGCCGTCGAAACTCCGTCATAGATAACCTCCGAGGAGGTGGTT	1416
DB	441	MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluValIle	460
QY	1417	GCAAAATTCATATGGCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1476
DB	461	GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnAsnProPhe	480
QY	1477	AAGTTCTTCCTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
DB	481	LysPheLeuValProGlnGluSerGlnLysArgAlaValAla	495

RESULT 4

PSYGI

glycinin chain A2Bla precursor - soybean

N;Alternate names: 11S globulin

N;Contains: glycinin chain A2; glycinin chain Bla

C;Species: Glycine max (soybean)

C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 09-Jul-2004

C;Accession: A91341; A92454; E92454; A90024; A92452; S71053; S74123; A05082; A05164; A051

R;Momma, T.; Negro, T.; Uda, K.; Fukazawa, C.

A;Title: A complete cDNA coding for the sequence of glycinin A2Bla subunit precursor.

A;Reference number: A91341

A;Accession: A91341

A;Molecule type: mRNA

A;Residues: 1-485 <MOM>

A;Cross-references: UNIPROT:P04405

A;Experimental source: strain Bonmimori

A;Note: the source of this mRNA was cotyledon tissue taken from seeds at the middle stage

R;Marco, Y.A.; Thanh, V.H.; Tumer, N.E.; Scallion, B.J.; Nielsen, N.C.

J. Biol. Chem. 259, 13436-13441, 1984

A;Title: Cloning and structural analysis of DNA encoding an A2Bla subunit of glycinin.

A;Reference number: A92454; MUID:85030472; PMID:6092376

A;Accession: A92454

A;Molecule type: mRNA

A;Residues: 262-446 <MA1>

A;Accession: B92454

A;Molecule type: DNA

A;Residues: 318-485 <MA2>

A;Experimental source: strain CX635-1-1-1

RESULT 10

S11003
glycinin G3 precursor - soybean
C:Species: Glycine max (soybean)
C;Date: 21-Nov-1993 #sequence_revision 19-Jan-1996 #text_change 03-May-1996
C;Accession: S11003
R;Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;
Plant Cell 1, 313-326, 1989
A;Title: Characterization of the glycinin gene family in soybean.
A;Reference number: S10851; MUID:92393391; PMID:2485233
A;Accession: S11003
A;Molecule type: DNA
A;Residues: 1-484 <NIE>
A;Experimental source: variety Dare
C;Genetics:
C;Gene: Gy3
C;Superfamily: glycinin
C;Keywords: storage protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-484/Product: glycinin G3 #status predicted <MAR>

Alignment Scores:
Pred. No.: 3.25e-98 Length: 484
Score: 1428.00 Matches: 290
Percent Similarity: 70.33% Conservative: 68
Best Local Similarity: 56.97% Mismatches: 95
Query Match: 53.07% Indels: 56
DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x S11003 (1-484)

Qy	1	CGG	CAG	CAAC	CGG	AGG	AGA	AGCG	GTCC	AGTTC	CAG	CGCT	CAAT	GTCC	GAG	AGAC	CTTG	AC	60			
Db	23	Arg	Glu	Gln	Pro	Gln	Asn	Glu	Cys	Gln	Ile	Gln	Arg	Leu	Asn	Ala	Leu	Lys	Pro	Asp	42	
Qy	61	AAT	CGC	ATT	CAAT	CAG	AGG	GGT	TAC	ATT	CAG	ACT	TGG	AAC	CCC	CAAC	ACC	AGG	AGT	TC	120	
Db	43	Asn	Arg	Ile	Glu	Ser	Glu	Gly	Gly	Phe	Ile	Glu	Thr	Trp	Asn	Pro	Asn	Asn	Lys	Pro	Phe	62
Qy	121	GAAT	CGC	CGG	CGG	CGC	CTC	CTC	CTG	CTT	AGT	CTC	CGC	CGG	CAAC	CGC	CTT	CGT	AGG	CCCT	180	
Db	63	Gln	Cys	Ala	Gly	Val	Ala	Leu	Ser	Arg	Cys	Thr	Leu	Asn	Arg	Asn	Ala	Leu	Arg	Arg	Pro	82
Qy	181	TTCT	ACT	CAAT	GTCC	CCG	AGG	AGAT	CTT	CAAT	CAG	CAAG	AGG	GGG	AT	ACT	TTT	GGG	TTG	CT	240	
Db	83	Ser	Tyr	Thr	Asn	Ala	Pro	Gln	Glu	Ile	Tyr	Ile	Gln	Gln	Gly	Ser	Gly	Ile	Phe	Gly	Met	102
Qy	241	ATAT	TCCT	GGT	GTG	TCCT	AGAC	ACT	ATG	AGAG	CGCT	CAC	ACA	CAAG	TCG	TCG	ATC	TC	AG	300		
Db	103	Ile	Phe	Pro	Gly	Cys	Pro	Ser	Thr	Phe	Glu	Glu	Pro	Gln	Lys	Gly	-----	Gln	119			
Qy	301	TCCC	AAAG	ACAC	CAAC	AGC	GTCT	CCA	AGG	AGA	GAG	CAAC	AAAG	CCAA	CAG	CAAC	AGAT	AGT	360			
Db	120	Ser	Ser	Arg	Pro	-----	-----	-----	-----	-----	-----	-----	-----	-----	Gln	Asp	Arg	126				
Qy	361	CACC	AGA	AGT	GTG	CAAC	CGT	TTT	CGA	GGT	GA	TCT	CAAT	TCAG	TTCC	CCG	GGT	TTG	CT	420		
Db	127	Hie	Gln	Lys	Ile	Tyr	His	Phe	Arg	Glu	Gly	Asp	Leu	Ile	Ala	Val	Pro	Thr	Gly	Phe	Ala	146
Qy	421	TTCT	GGCT	CTC	TAC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	AGC	480		
Db	147	Tyr	Trp	Met	Tyr	Asn	Asn	Glu	Asp	Thr	Pro	Val	Val	Ala	Val	Ser	Leu	Ile	Asp	Thr	Asn	166
Qy	481	AACA	ACC	GAC	CAAC	CAAC	AGT	TCAG	TTCC	CCG	AGG	AGAT	TTCA	ATT	TGG	CTG	GGA	CAAC	CGG	AG	540	
Db	167	Ser	Phe	Gln	Asn	Gln	Leu	Asp	Gln	Met	Pro	Arg	Arg	Phe	Tyr	Leu	Ala	Gly	Asn	Gln	Glu	186
Qy	541	CAAG	AGT	TTT	TAAG	GTAC	CAAC	CAAC	AAAG	CAG	CAAC	AAAG	CAG	CAAC	AAAG	CAG	CAAC	AAAG	CAG	CAAC	600	
Db	187	Gln	Glu	Phe	Leu	Gln	Tyr	Gln	Pro	Gln	Lys	Gln	Gly	-----	-----	-----	-----	-----	-----	199		
Qy	601	AGCC	CAT	ACAG	CCCG	CAAG	TGCT	CAG	CTT	AG	CAAC	AGAG	CGT	GAAT	TTAG	CCCT	TCG	AGGA	660			

RESULT 11

S14393
legumin A2 precursor - fava bean
C;Species: Vicia faba (fava bean)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S14393
R;Schlesier, B.; Bassuener, R.; van Hai, N.; Muentz, K.
Nucleic Acids Res. 18, 7146, 1990
A;Title: The cDNA derived primary structure of two distinct legumin A subunit precursors
A;Reference number: S14392; MUID:91088307; PMID:2263481

A;Accession: S14393
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-500 <SCH>
A;Cross-references: UNIPROT:Q99304; EMBL:X55014; NID:g222007; PIDN:CAA38758.1; PID:g22008
C;Superfamily: glycinin

Alignment Scores:

Pred. No.: 3 29e-97 Length: 500
Score: 1414.50 Matches: 280
Percent Similarity: 67.30% Conservative: 76
Best Local Similarity: 52.93% Mismatches: 100
Query Match: 52.56% Indels: 73
DB: 2 Gaps: 8

US-10-728-323-3 (1-1524) x S14393 (1-500)

```

QY 1 CGGCAGAACCGGAGGAGAACGGTCCAGTTCACGGCTCAATCGCGAGAGACCTGAC 60
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
QY 61 AATCGCATTGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 43 AsnArgIleGlnSerGluGlyGlyLeuIleGluThrIleAsnProAsnAsnArgGlnPhe 62
QY 121 GAATGGCGCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 63 ArgCysAlaSerValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCATGCTCCAGGAGATCTTCATCCAGCAAGAAAGGGATCTTTGGGTG 240
Db 83 TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet 102
QY 241 ATATTCCTCGTGTCTGACACTATGAAGCCCTCACACAAAGTGTGCTGATCTCAG 300
Db 103 ValPheProSerCysProGluThrPheGluGluPro-----GlnGln 116
QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGAGACCAAGGCCAACAGCAACGAGATAGT 360
Db 117 SerGlu-----GlnGlyGlu-----GlyGlyArgTyrArgAspSer 128
QY 361 CACCAGAGTGCACCGTTTCGATAGGAGGTGATCTCATTCGAGTTCACCGGTGTGCT 420
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148
QY 421 TTCTGCTCTACACACGACGACACTGATGTTGTGCTGTTCTCTTACTGACACCAAC 480
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspIleGly 168
QY 481 AACACGACACACCGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAACACGGAG 540
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
QY 541 CAAGAGTCTTAAGGTACGACCAACAAAGCAGACAAAGCAGACGAGAAAGAGCTTACCATAT 600
Db 189 GlnGluPheLeuArgTyrGlnHisGln----- 197
QY 601 AGCCCATACGCCGCAAGTCCAGCTCAGCTAGACAGAGAGCGTGAATTTAGCCCTCAGGA 660
Db 197 ----- 197
QY 661 CAGCAGACCGCAGAGAACGAGCAGGACCAAGAGAAAGAAACGAAAGGTGGAAACATCTTC 720
Db 198 -----GlnGlyGlyGlyGluGluGlnAspAsnAspGlyAsnAsnIlePhe 212
QY 721 AGCGGCTTACCGCGGAGTTCCTGGAAACAGCCCTCCAGTTGACACGACAGACATAGTG 780
Db 213 SerGlyPheIleArgAspPheLeuGluAspAlaLeuAsnVal---AsnArgHisIleVal 231
QY 781 CAAACCTTAAGAGCGCAGACCGGAGCTGAAGAGAGGAGCCATTCTGACAGTGGAGGA 840
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGlyGlyAlaIleValIleValIleValIleGly 921

```

```

QY 841 GGCCTCAGAACTCTTGAGCCAGCCAGATAGAAAG-----AGACGTCCCGACGAA 885
Db 252 GlyLeuSerIleIleThrProGluArgGlnAlaArgHisProArgGlySerArgGln 271
QY 886 GAAAGGAATACGATCAGATCAATATGAA----- 915
Db 272 GluGluAspGluAspGluAspGluGlyGluArgGlnProSerHisIleSerArg 291
QY 916 -----TACGATGAAGAGGATAGA-----AGCGGTGGCAGGGAAGCAGA 954
Db 292 ArgGlyGluAspGluAspAspGlyGluLysArgHisSerGlnLysGlyGluSerArgArg 311
QY 955 GGCAGGGGAATGGTATTGAAGAGACGATCTCCACCGCAAGTGCTAAAGAACATTTGT 1014
Db 312 HisGlyAspAsnGlyLeuGluGluThrValCysThrAlaLysLeuAsnIleGly 331
QY 1015 AGAACACAGATCCCTGACATCTACACCTCAAGCTGGTTCACTCAAACTGCCAACGAT 1074
Db 332 SerSerSerSerProAspIleTyrAsnProGlnAlaGlyArgIleIleThrValThrSer 351
QY 1075 CTCACCTTCTAATCTAGTGGCTTGACCTAGTCTGTAATATGGAATCTCTACAGG 1134
Db 352 LeuAspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLys 371
QY 1135 AATGCAATTTGTTCGCTCCTACACCAACCAACGACACAGCATCATATATCGATTGAGG 1194
Db 372 AsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleLeuTyrAlaLeuLys 391
QY 1195 GGACGGCTCAGCTCGAAGTCTGACACGACGACGACGACGAGTGTACGACGAGGCTT 1254
Db 392 GlyArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGluGluLeu 411
QY 1255 CAAGAGGGTCAGCTGCTGTGTCGCACACAGAACTTCGCGCTGCGTGGAAAGTCCACAG 1314
Db 412 GluAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSer 431
QY 1315 GAGAACTTCGAATCAGTGCATTCGAAGCAGACATCAAGCCCGCAGCATGCCAACCTCGCC 1374
Db 432 AspArgPheThrTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAla 451
QY 1375 GGTGAAATCCGTCATAGATAACCTCGCGGAGGAGTGGTTGCCAAATTCATATGGCCTC 1434
Db 452 GlyThrSerSerValIleAsnAspMetProValAspValValAlaAlaThrPheAsnLeu 471
QY 1435 CAAAGGAGCAGCAGGAGGCTTAAAGAACAAACCAACCCCTTCAAGTTCCTCGTCCACCG 1494
Db 472 GluArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProPro 491
QY 1495 TCTCAGCAGTCTCCGAGGCTGTGCT 1521
Db 492 ArgGluSerGlnLysArgAlaSerAla 500

```

RESULT 12

T06452
probable legumin A precursor - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T06452
R:Lycett, G.W.; Croy, R.R.D.; Shirsat, A.H.; Richards, D.M.; Boulter, D.
Nucleic Acids Res. 13, 6733-6743, 1985
A:Title: The 5'-flanking regions of three legumin genes: comparison of the DNA sequences.
A:Reference number: Z15687; MUID:86041868; PMID:2997721
A:Accession: T06452
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-507 <LYC>
A:Cross-references: EMBL:X02982; NID:g20777; PIDN:CAA26720.1; PID:g4379378
C:Genetics:
A:Introns: 96/1; 179/3; 388/3
C:Superfamily: glycinin
C:Keywords: seed; storage protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-332/Product: legumin A, alpha chain #status predicted <ALP>

Alignment Scores:
 Pred. No.: 6.1e-95 Length: 507
 Score: 1384.00 Matches: 283
 Percent Similarity: 63.92% Conservative: 66
 Best Local Similarity: 51.83% Mismatches: 97
 Query Match: 51.43% Indels: 100
 DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x T06452 (1-507)

```

QY 1 CGGCAGCAACCGGAGAGAACCGCTGCCAGTTCAGCGCCTCAATGGCAGAGACCTGAC 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42

QY 61 AATCGCATTAATCAGAGCGGTTACATTGAGACTTGGAAACCCCAACACCGAGGATTC 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 43 AsnArgileGluSerGluGlyGlyLeuileGluThrTrpAsnProAsnAsnLysGlnPhe 62

QY 121 GAATGCGCGCGCTCGCCCTCTCGCTTAGTCTCGCGCGCAACCGCCTTCGTAGCCT 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro 82

QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGAAAGGGGATACTTTGGGTTG 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 TyrTyrSerAsnAlaProGlnGluilePheileGlnGlyAsnGlyTyrPheGlyMet 102

QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAGGTCTCGATCTCAG 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 ValPheProGlyCysProGluThrPheGluPro-----GlnGlu 116

QY 301 TCCCAAGAACCCAGACGCTCCCAAGGAGAGAACCAAGCCACAGCAACGAGATAGT 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspArg 128

QY 361 CACGAGAGGTGCACCGTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGTGTGCT 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleleAlaValProThrGlyIleVal 148

QY 421 TTCGTGCTCTACACGACACGACACTGATGTTGTGCTGTTCTCTTACTGACACCAAC 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg 168

QY 481 AACACGACACACGCTTGATGATTCCTCCAGAGATTCATTTGGTGGGAACACGGAG 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu 188

QY 541 CAAGAGTTCTTAAGGTACACGACAAAGCAGACAAAGCAGACGAGAAAGCTTACCATAT 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 GlnGluPheLeuGlnTyrGlnHisGln----- 197

QY 601 AGCCCATACAGCCCGCAAAAGTACGCTTAGACAAGAGCGTGAAATTTAGCCCTCGAGGA 660
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 197 ----- 197

QY 661 CAGCAGACCGCGAGAGAACGAGCAGGACAGAGAACAGAAACAGAGTGAACATCTTTC 720
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 -----GlnGlyGlyLysGlnGluGlnGluAsnGluGlyAsnAsnIlePhe 212

QY 721 AGCGGCTTACGCGCGGAGTTCTCGGACAGACCTTCAGGTTCCAGCAGACAGATAGTG 780
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 SerGlyPheLysArgAspTyrLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 231

QY 781 CAAAACTTAAGCGGAGACCGGAGAGTGAAGAGAGGGAGCCATTGTGACAGTGAGGGGA 840
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGluGlyAlaIleValLysValLysGly 251

QY 841 GGCTCTAGATCTTGAGCCCGCAGATAGAAG----- 870
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 GlyLeuSerIleleSerProGluLysGlnAlaArgHisGlnArgGlySerArgGln 271

QY 871 -----AGCGTCCGAGGAAGAG 891
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 272 GluGluAspGluAspGluLysGlnProArgHisGlnArgGlySerArgGlnGluGlu 291
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGGATAGAGCGTGCC----- 942
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 GluGluAspGluAspGluGluArgGlnProArgHisGlnArgArgGlyGlyGluGlu 311
QY 943 -----AGGGGAACGAGGAGCGGGG----- 963
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 GluGluAspLysLysGluArgGlyGlySerGlnLysGlyLysSerArgArgGlnGlyAsp 331
QY 964 AATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAAAAGAACATTTGGTAGAACAGA 1023
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 332 AsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlyProSerSer 351
QY 1024 TCCCTGACATCTACAACTCAAGCTGTGTTCACTCAAACTGCCAACGATCTCAACCTT 1083
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 352 SerProAspIleTyrAsnProGluAlaGlyArgIleLysThrValThrSerLeuAspLeu 371
QY 1084 CTAATACTTAGTGGCTTCGACCTAGTCTGATATGGAATCTCTACAGGAATGCATTG 1143
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 ProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuHisLys----- 388
QY 1144 TTGTGCTCACTACAAACCAACGACACGACATCATATATCGATTGAGGGACCGGCT 1203
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 389 -----AsnAlaAsnSerIleIleTyrAlaLeuLysGlyArgAla 401
QY 1204 CACGTGCAAGTCTGTGACAGCAACGCGACAGAGTGTACGAGAGGATTCGAAGGGT 1263
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 402 ArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGlyLeuGluAlaGly 421
QY 1264 CACGTGCTGTGTGTCACAGAACCTTCGCGCTCGTGGAAAGTCCCGAGGAGACTTC 1323
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 ArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSerAspArgPhe 441
QY 1324 GAATAGTGGCATTTCAAGACAGACTCAAGGCCAGCATACGCAACCTCGCGGTGAAAC 1383
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 442 SerTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSer 461
QY 1384 TCCGTATAGATAAAGTCCGAGGAGGTGGTTCAAATTCATATATGCGCTCCAAAGGAG 1443
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 462 SerValIleAsnAsnLeuProLeuAspValValAlaAlaThrPheAsnLeuGlnArgAsn 481
QY 1444 CAGGCAAGGACCTTAAGAACACACACCTTCAGTTCCTTCGTTCCCGCTCCAGCAG 1503
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProAlaArgGluSer 501
QY 1504 TCTCCGAGGCTGTGGCT 1521
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 502 GluAsnArgAlaSerAla 507

RESULT 13
S14392
legumin A1 precursor - fava bean (fragment)
C:Species: Vicia faba (fava bean)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S14392
R:Schlesier, B.; Baasuenex, R.; van Hai, N.; Muentz, K.
Nucleic Acids Res. 18, 7146, 1990
A:Title: The cDNA derived primary structure of two distinct legumin A subunit precursors
A:Reference number: S14392; MUID:91088307; PMID:2263481
A:Accession: S14392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-497 <SCH>
A:Cross-references: UNIPROT:Q03971; EMBL:X55013; NID:g22006; PIDN:CAA38757.1; PID:g388251
C:Superfamily: glycinin

Alignment Scores:
Pred. No.: 2.84e-94 Length: 497
Score: 1375.00 Matches: 276
Percent Similarity: 68.80% Conservative: 79
Best Local Similarity: 53.49% Mismatches: 107
Query Match: 51.10% Indels: 54

```


DB: 2 Gaps: 11
US-10-728-323-3 (1-1524) x S14392 (1-497)
QY 25 TGCAGTTCAGGCGCTCAATGCGCAGAGACCTGACAAATCGCATTTGAATCAGAGGGCGGT 84
D 1 CysGlnLeuGluArgLeuAspAlaLeuGluProAspAsnArgIleGluSerGluGly 20
QY 85 TACATTGAGACTTGGAAACCCCAACACAGGAGTTGGAATGCGCCGCGTCCCTCTCT 144
D 21 LeuIleGluThrIrpAsnProAsnAsnArgGlnPheArgCysAlaArgValAlaLeuSer 40
QY 145 CGTTAGTCTCCGCCCGCAACGCGCTTCGTAGGCGTTCTTACTCCAATGCTCCCCAGGAG 204
D 41 ArgAlaThrLeuGlnArgAsnAlaLeuArgArgProIrfYrSerAsnAlaProGlnGlu 60
QY 205 ATCTTCATCCAGCAGAGGAGATCTTTGGGTGGTGGATATTCCTGGTGTCTCTAGACAC 264
D 61 IleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMetValPheProSerCysProGluThr 80
QY 265 TATGAAGAGCCTCACACACAAGGTCTGTCGATCTCAGTCCCAAGACACCAAGACGCTCTC 324
D 81 PheGlnGluPro-----GlnGlnSerGlu----- 88
QY 325 CAAGGAGAACCAAAAGCCCAACACGACGATAGTACACAGAGGTGCACCGCTTTCGAT 384
D 89 GlnGlyGlu-----GlyArgArgTyrArgAspSerHisGlnLysValAsnArgPheArg 106
QY 385 GAGGTGATCTCATGTCAGTTCACCGGTTGCTTCTTCTGCTCTTACAAACGACACGAC 444
D 107 GlnGlyAspIleIleAlaValProThrGlyIleValPheTrpMetTyrAsnAspGlnAsp 126
QY 445 ACTGATGTTCTGCTGTTCTCTACTGACACCAACAAACACACACAGCTTGTATCAG 504
D 127 IleProValIleAlaIleSerLeuThrAspThrGlySerSerAsnAsnGlnLeuAspGln 146
QY 505 TTCCCGCAGGAGATTCAATTTGGCTGGGAAACGCGAGCAAGAGTCTTAAAGGTACCAACAA 564
D 147 MetProArgArgPheTyrLeuAlaGlyAsnGlnGluGlnPheLeuArgTyrGlnHis 166
QY 565 CAAAGCAGACAAGCAGACAGAGCTTACCATATAGCCCATACAGCCGCGAAAGTCTAG 624
D 167 GlnGlnGlyValLysGluGlu-----GlnAspAsn 176
QY 625 CCTAGACAAGACAGCGTGAATTTAGCCCTCGAGGACAGCACAGCCGACAGAAACGACCA 684
D 177 AspGlyAsnGlnGlnGluPheLeu---ArgTyrGlnHis-----ArgGlnGlyVal 193
QY 685 GGACAAGAAGAAACGAAAGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCCTG 744
D 194 LysGluGlnAspAsnAspGlyAsnAsnIlePheSerGlyPheAsnArgAspPheLeu 213
QY 745 GAACACGCTTCAGGTTGACGACAGACAGATAGTCAAAACCTAAGAGCGGACCGAG 804
D 214 GluAspAlaPheAsnVal---AsnArgHisIleValAspArgLeuGlnGly-----Arg 230
QY 805 AGTGAAGAAGAGGAGCCATTGTGACAGTGGGAGGCGCTCAGAAATCTTGAGCCCCAGAT 864
D 231 AsnGlnGluArgGlyAlaIleValLysValLysGlyGlyLeuSerIleIleThrProPro 250
QY 865 AGAAAGAGA-----CGTCCCGCAGCAAGAGGAATACGATGAA 903
D 251 GluArgGlnAlaArgHisProArgGlySerArgGlnGluAspGluAspGlu 270
QY 904 GATGAATGAA-----TACGATGAAGAG 927
D 271 AspGlnLysGluArgGlnProSerHisHisLysSerArgArgGlyGluAspGluAsp 290
QY 928 GATAGA-----AGCGGTGGCGAGGAGCAGGCGAGGGGAATGTATT 972
D 291 AspLysGlnLysArgHisSerGlnLysGlyLysSerArgArgHisGlyAspAsnGlyLeu 310
QY 973 GAAGAGCGATCTGCACCGCAAGTCTGCTAAAGAAACATTGGTAGAAACAGATCCCTCGAC 1032

Db 311 GluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlySerSerSerProAsp 330
QY 1033 ATCTACACCCCTCAAGCTGGTTCACCTCAAACTGCCAAGCATCTCAACCTTCTAATACTT 1092
D 331 IleTyrAsnProGlnAlaGlyArgIleYsThrValThrSerValAspLeuProValLeu 350
QY 1093 AGGTGGCTGGACCTAGTGTGAATATGAAATCTCTACAGGAATGCATTTGTTGCTGCT 1152
D 351 ArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLysAsnAlaMetPheValPro 370
QY 1153 CACTACAAACCAACGACACAGCATCATATATCATGATTGAGGGACGGGCTCACGTGCAA 1212
D 371 HisTyrAsnLeuAsnAlaAsnSerValLeuTyrAlaLeuLysGlyArgAlaArgLeuGln 390
QY 1213 CTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTT 1272
D 391 ValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaGlyArgAlaLeu 410
QY 1273 GTGGTGCCACAGAACTTCGCCGCTGGAAAGTCCAGAGCGAGAACTTCGAATACGTG 1332
D 411 ThrValProGlnAsnTyrValValAlaAlaLysSerLeuSerAspArgPheThrTyrVal 430
QY 1333 GCATTCAAGACAGACTCAAGCCCATAGCAACCTCGCCGTCGAAACTCCGTCATA 1392
D 431 AlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSerSerValIle 450
QY 1393 GATAACCTCCCGAGGAGGTGTTGCAATTCATATGCTCCAAAGGAGGAGCAGCGAAG 1452
D 451 AsnAspLeuProLeuAspValValAlaAlaThrPheAsnLeuGluArgAsnGluAlaArg 470
QY 1453 CAGCTTAAAGAAACAAACCCCTTCAAGTTCCTTCGTTCCACCGCTCTCAG 1500
D 471 GlnLeuLysPheAsnAsnProSerArgPheLeuValProArgGlu 486
RESULT 14
S51941
N:Alternate names: salt-soluble globulin; seed storage protein
C:Species: Prunus dulcis (almond)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C:Accession: S51941; S42474
R:Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.
Plant Mol. Biol. 27, 205-210, 1995
A:Title: Molecular characterization of cDNAs corresponding to genes expressed during alme
A:Reference number: S51940; MUID:95170003; PMID:7865791
A:Accession: S51941
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-551 <GAR>
A:Cross-references: UNIPROT:Q43607; EMBL:X78119
A>Note: the source is designated as Prunus amygdalus
R:Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.
submitted to the EMBL Data Library, March 1994
A:Description: Molecular characterization of cDNAs corresponding to proteins related to e
A:Reference number: S42473
A:Accession: S42474
A:Molecule type: mRNA
A:Residues: 1-60, 'G', 62-551 <GAR>
A:Cross-references: EMBL:X78119; NID:G460805; PIDN:CAA55009.1; PID:G460806
A>Note: the source is designated as Prunus amygdalus
C:Superfamily: glycinin
C:Keywords: seed; storage protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-367/Product: prunin 1 alpha chain #status predicted <MAT1>
F:368-551/Product: prunin 1 beta chain #status predicted <MAT2>
F:108-374/Disulfide bonds: #status predicted
Alignment Scores:
Pred. No.: 7 47e-72 Length: 551
Score: 1073.50 Matches: 228
Percent Similarity: 57.22% Conservative: 89
Best Local Similarity: 41.16% Mismatches: 160

US-10-728-323-3 (1-1524) x S11004 (1-560)

QY	19	AACCGCTGCAGTTCCAGCGCCTCAATGCGCAGAGACCTGACAACTGCATTGTAATCAGAG	78
Db			
Db	30	AsnGluCysGlnLeuAsnAsnLeuAsnAlaLeuGluProAspHisArgValGluSerGlu	49
QY	79	GGCGGTTCATTGAGACTTTGGAACCCCAACCAACAGAGGTTTCGAATCGCGCGCGTCGCC	138
Db			
Db	50	GlyGlyLeuIleGlnThrTrrAsnSerGlnHisProGluLeuLysCysAlaGlyValThr	69
QY	139	CTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCTTCTTACTCAATGCTCC	198
Db			
Db	70	ValSerLysLeuThrLeuAsnArgAsnGlyLeuHisLeuProSerTyrSerProTyrPro	89
QY	199	CAGGAGATCTTCATCCAGCAAGAAAGGGATACTTTGGTTTGATATATTCCTCGTGTCTCT	258
Db		:::	
Db	90	ArgMetIleIleIleAlaGlnGlyGlyAlaLeuGlyValAlaIleProGlyCysPro	109
QY	259	AGACACTATGAAGAGCTCACACAGGTCGTGATCTGTCGCCAAAGACCAACCAAGA	318
Db		:::	
Db	110	GluThrPheGluGluProGlnGlnSerAsnArgArgGlySerArg	125
QY	319	CGTCTCAAGGAGAACACCAAGGCCAACAGCAACAGATAGTCCACCAAGAGGTGCACCGT	378
Db			
Db	126	-----SerGlnLysGlnGlnLeuGlnAspSerHisGlnLysIleArgHis	140
QY	379	TTCCGATGAGGGTGATCTCATTCGACTCCCAACCGGTGTGCTTCTGGCTCTCAACAGAC	438
Db			
Db	141	PheAsnGluGlyAspValLeuValIleProGlyValProTyrTrrThrTyrAsnThr	160
QY	439	CAGCACACTGATGTGTGCTTCTCTTACTGACACCAACAGCAACAGCAACACAGCTT	498
Db			
Db	161	GlyAspGluProValValAlaIleSerLeuLeuAspThrSerAsnPheAsnAsnGlnLeu	180
QY	499	GATCAGTCTCCGAGGAGATTCAATTTGGCTCGGAACACGGAGCAAGAGTCTTTAAGGTAC	558
Db			
Db	181	AspGlnThrProArgValPheTyrLeuAlaGlyAsnProAspIleGluTyr	197
QY	559	CAGCAACAAGCAGACNAAGCAGACGAGAGAGCTTACCATATAGCCCATACAGCCCGCAA	618
Db			
Db	198	-----ProGluThrMetGln	202
QY	619	AGTCAGCTTAGACAAGAAGAGCGGTCAATTTAGCCCTCGAGCAGACAGCCGACAGAA	678
Db			
Db	203	GlnGlnGlnGlnLysSerHisGlyGlyArgLysGlnGlyGlnHisGln	219
QY	679	CGAGCAGGACAGAGAGAAACCAAGAGTGGAAACATCTTCAGCGGCTTCACCGCGGAG	738
Db			
Db	220	-----GlnGluGluGluGluGlyGlySerValLeuSerGlyPheSerLysHis	236
QY	739	TTCTCTGGAAACAGCCTTCCAGGTTCCACACAGACAGATAGTGTAAAAACCTTAAGAGCGGAG	798
Db			
Db	237	PheLeuAlaGlnSerPheAsnThrAsnGlu---AspIleAlaGluLysLeu	253
QY	799	ACCGAGGTGAAGAAGAGGAGCCATTGTGACGTGAGGGAGGCGCTCAGAATCTTTGAGC	858
Db			
Db	254	SerProAspAspGluArgLysGlnIleValThrValGluGlyGlyLeuSerValIleSer	273
QY	859	CCAGATAGAAGAGAGCGTCCGACCAAGACGAGGATACATGAA	903
Db			
Db	274	ProLysTrrPgnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu	293
QY	904	-----GATGAATAT	912
Db			
Db	294	GlnIleProSerHisProProArgProSerHisGlyLysArgGluGlnAspGluAsp	313
QY	913	GAATACGATGAAGAGATAGA---AGGCGTGGCAGGGAGACGAGCG	957
Db			
Db	314	GluAspGluAspGluAspGluProSerArgProSerGlnGlnLysArgGluGln	333
QY	957	-----	957

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 93.2661 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10728323/runat 23082005 124355 29210/app_query.fasta 1.4757
-DB=Uniprot 03 -Qfmt=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DUALIGN=200 -THR_SCORE=pept -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1 291 @runat 23082005 124355 29210 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPOP=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2673	99.3	507	2	O82580
2	2479	92.1	529	2	O9FZ11
3	2455	91.2	536	2	O647H4
4	2451.5	91.1	537	2	O647H3
5	2444	90.8	538	2	O8LKN1
6	2436	90.5	536	2	O6T2T4
7	2421	90.0	530	2	O9SQH7
8	1834	68.2	510	2	O61WG5
9	1651	61.4	484	2	O647H2
10	1502	55.8	481	1	GLC3_SOYBN
11	1499.5	55.7	485	1	GLC2_SOYBN
12	1496	55.6	495	1	GLC1_SOYBN
13	1489	55.3	481	2	O852U5
14	1488.5	55.3	482	2	O852U4
15	1457.5	54.2	498	2	O41702
16	1455	54.1	520	1	LEG2_PEA

17	1443	53.6	517	2	O9TOP5
18	1440	53.5	517	1	LEG4_PEA
19	1431.5	53.2	482	2	O41676
20	1414.5	52.6	500	2	O99304
21	1375	51.1	497	2	O03971
22	1186	44.1	496	2	O9SMJ4
23	1110.5	41.3	515	2	O8W1C2
24	1074.5	39.9	488	2	O41128
25	1069.5	39.7	551	2	O43607
26	1059	39.4	542	2	O8LK20
27	1049	39.0	560	2	O9S9D0
28	1034.5	38.4	457	2	O8GZP6
29	1032	38.4	503	1	LEGJ_PEA
30	999.5	37.1	566	2	O24234
31	998	37.1	485	2	O41703
32	996.5	37.0	484	1	LEG4_VICFA
33	986	36.6	517	2	O39922
34	986	36.6	517	2	O7GC77
35	985	36.6	517	2	O9SB12
36	981.5	36.5	564	2	O43673
37	980	36.4	517	2	P93707
38	980	36.4	517	2	P93708
39	978.5	36.4	219	2	O8LL03
40	975	36.2	563	2	O9SB11
41	974	36.2	563	2	O39921
42	973.5	36.2	562	1	GLC4_SOYBN
43	968	36.0	504	2	O43608
44	951.5	35.4	662	2	O647H1
45	946.5	35.2	562	2	O43452

ALIGNMENTS

RESULT 1

O82580 PRELIMINARY; PRT; 507 AA.

AC O82580; O82580; (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Glycinin (Fragment).
GN Name=Arah3;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99146968; PubMed=10021462;
RA Rabjohn P., Helm E.M., Stanley J.S., West C.M., Sampson H.A.,
RA Burks A.W., Bannan G.A.;
RT "Molecular cloning and epitope analysis of the peanut allergen Ara h
RT 3.";
RL J. Clin. Invest. 103:535-542(1999).
DR EMBL; AF093541; AAC63045.1; -.
DR HSSP; P04776; IUD1.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR Q8LKN1; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_lls.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 507 AA; 58349 MW; F3FB38BC3CB82DED CRC64;

Alignment Scores:
Pred. No.: 5,12e-179 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 99.33%		Indels: 0
DB: 2		Gaps: 0
US-10-728-323-3 (1-1524) x 082580 (1-507)		
QY	1	CGGAGCAACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTGAC 60
DB	1	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProasp 20
QY	61	AATCGCATTAATCAGAGGGCGGTTCATTTGAGACTTGGAAACCCCAACACAGAGGATTC 120
DB	21	AsnArgileGluSerGluGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 40
QY	121	GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGGCCT 180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAenAlaLeuArgArgPro 60
QY	181	TTCTACTCCAATGCTCCCCAGGAGATTTTCATCCAGCAAGAAAGGGGATATTTGGGTTG 240
DB	61	PheTyrSerAenAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY	241	ATATTCCCTGGTGTCTAGACTATGAAGAGCCTCACACAAAGTGTGTCATCTCAG 300
DB	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 100
QY	301	TCCCAAGACCAACAGACGTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
DB	101	SerGlnArgProProArgArgLeuGlnGlyIleGlnGlnGlnGlnGlnGlnGlnGlnGln 120
QY	361	CACCAGAAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCTCCACCGGTGTGCT 420
DB	121	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY	421	TTCTGGCTTACAAACGACGACGACTGATGTTGCTGCTTCTTCTTACTGACACCAAC 480
DB	141	PheTrpLeuTyrAenAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY	481	AACAACGACCAACGAGTTCATGATTCCTCCAGGAGATTCATTTGGCTGGGAACGGAG 540
DB	161	AsnAsnAspGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY	541	CAAGAGTCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
DB	181	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200
QY	601	AGCCCATACAGCCGCAAGTTCAGCTAGACAGAGAGCGTGAATTTAGCCCTCCAGGA 660
DB	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY	661	CAGCAGACCGCAGAGAACGAGCAGGACAGACAGAAAGAAACCAAGGTGGAACATCTTC 720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluGluGluGluGluGluGluGly 240
QY	721	AGCGGTTCACGCGGAGTTCCTGGAACAAGCCTTCAGGTTCCAGCAGACAGACAGATAGTG 780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal 260
QY	781	CAAACTAGAGCGGAGCCGAGAGTGAAGAGAGGAGCCATTGTGACAGTACGGGA 840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGly 280
QY	841	GGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCTGCCGACCAAGAAAGAGGAATACGAT 900
DB	281	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 300
QY	901	GAAAGTGAATATGAATACGATGAAGAGGATAGAGGCGGTGGCAGCGGAACGAGCGCAG 960
DB	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY	961	GGGAATGATTTGAGAGAGCATCTCCACCGCAAGTCTTAAAGAACATTGTTAGAAAC 1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340

RESULT 2

ID	Q9FZ11	PRELIMINARY;	PRT;	529 AA.
AC	Q9FZ11;			
DC	01-MAR-2001 (TReMBLrel. 16, Created)			
DT	01-WAR-2001 (TReMBLrel. 16, Last sequence update)			
DT	01-WAR-2004 (TReMBLrel. 26, Last annotation update)			
DE	Gly1.			
GN	Name=Gly1;			
OS	Arachis hypogaea (Peanut).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;			
OC	Arachis.			
OX	NCBI_TaxID=3818;			
RN	[1]_TaxID=3818;			
RP	SEQUENCE FROM N.A.			
RA	Jain A.K., Basha S.M.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF125192; AAG01363.1; -			
DR	HSSP; P04776; 1UCX			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR006045; Cupin.			
DR	InterPro; IPR007113; Cupin region.			
DR	InterPro; IPR011051; RmC like cupin.			
DR	Pfam; PF00190; Cupin; 2			
DR	PRINTS; PR00439; 11SGLOBULIN.			
DR	PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN 1.			
SQ	SEQUENCE 529 AA; 60448 MW; BB6F25BC1D6E06A1 CRC64;			
Alignment Scores:				
Pred. No.:	2,11e-165	Length:	529	
Score:	2479.00	Matches:	475	
Percent Similarity:	95.28%	Conservative:	9	
Best Local Similarity:	93.50%	Mismatches:	20	

Query Match:	92.12%	Indels:	4
DB:	2	Gaps:	2
US-10-728-323-3 (1-1524) x Q9F211 (1-529)			
QY	1	CGGCAGCAACCGGAGGAGAACGGCTGCAGTTCACGCGCTCAATGCGCAGACCTGAC	60
DB	25	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnA-gProasp	44
QY	61	AATGCGATTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACAAACCGAGGATTC	120
DB	45	AsnArgLeuGluSerGluGlyGlyTyrlleGluThrPheAsnProAsnAsnGlnGluPhe	64
QY	121	GAATGCGCGCGCTGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCGCTTCGTAGSCCT	180
DB	65	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	84
QY	181	TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTTG	240
DB	85	PheTyrsSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrllePheGlyLeu	104
QY	241	ATATTCCTGTTGCTCTAGACACTATGAAGAGCTCACACAAAGTTCGTCTGATCTCAG	300
DB	105	IlePheProGlyCysProSerThrTyrlleGluProAlaGlnGlnGlyArgGlnHisGln	124
QY	301	TCCCAAGACACCAAGAGCTTCCAAAGGAGAGCAACCAAGC---CAACAGCAACGAGAT	357
DB	125	SerGlnArgAlaProArgArgPheGluGlyCysGlnSerGlnGlnGlnGlnGlnAsp	144
QY	358	AGTCACAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCCTCCACCGGTGT	417
DB	145	SerHisGlnLysValArgA-gPheAspGluGlyAspLeuIleAlaValProThrGlyVal	164
QY	418	GCTTTCTGCTCTACACGACCAACGACACTGATGTTGTTGCTGTTCTTCTTACTGACAC	477
DB	165	AlaLeuTrpMetPheAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr	184
QY	478	AACAAACACCAACACAGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACAG	537
DB	185	AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHis	204
QY	538	GAGCAAGAGTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACCAAGAGCTTACCA	597
DB	205	GluGlnGluPheLeuArgTyrlleGlnGln-----GlnSerArgArgSerLeuPro	221
QY	598	TATAGCCCATACAGCCCGCAAGTGCAGCTAGACAAGAGCGTGAATTTAGCCCTCGA	657
DB	222	TyrlleSerProTyrlleSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg	241
QY	658	GGACAGCACCGCCAGAGAACGAGCAGGACAGAAAGAAAGAAAGAGGTGGAAACATC	717
DB	242	GlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIle	261
QY	718	TTACAGCGCTTCAGCGCGAGTTCCTCGAAACAAGCTTCAGAGTTGACGACAGACAGATA	777
DB	262	PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspArgGlnIle	281
QY	778	GTGCAAAACCTTAAGAGCGCAGACCGAGAGTCAAGAGAGGGAGCCATTGTGACAGTGAGG	837
DB	282	ValGlnAsnLeuArgGlyGluAsnGluSerGluGluGlyAlaIleValThrValLys	301
QY	838	GGAGGCTCAGATCTTTGAGCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAATAC	897
DB	302	GlyGlyLeuArgIleLeuSerProAspArgLysArgGlyAlaAspGluGluGluGluTyrlle	321
QY	898	GATCAAGCATGAATGAATACGATGAAGAGATAGAGGCTGCGAGGGGAACGAGGC	957
DB	322	AspGlnAspGluTyrlleGluTyrlleAspGluLysArgArgArgGlyArgGlySerArgGly	341
QY	958	AGGGGGAATGGATTGAAGAGACCATCTGCAACCGCAAGTGTCAAAAGAACATTGGTAGA	1017
DB	342	ArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerValLysLysAsnIleGlyArg	361
QY	1018	AACAGATCCCTGACATCTACACCTCAAGCTGGTTCACCTCAAAACTGCCACGATCTC	1077
DB	362	AsnArgSerProAspIleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeu	381
QY	1078	AACCTTTCTAATACTAGGTGGCTTGACCTAGTGTGCTGAATATGAAATCTCTACAGGAAT	1137
DB	382	AsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrlleAsnLeuTyrlleArgAsn	401
QY	1138	GCATTTGTTGCTCCTCACTACAAACCAACGACACAGCATCATATATATGATTCAGGGGA	1197
DB	402	AlaLeuPheValProHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleAlaLeuArgGly	421
QY	1198	CGGGCTCAGTGCAGTTCGTCGACCAACGCGCAACAGAGTGTCAGCAGGAGGCTTCAA	1257
DB	422	ArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluLeuGln	441
QY	1258	GAGGCTCAGTTCGTCGTCGACCAAGAACTTCGCGCTGCTCGAAAGTCCACGAGCGAG	1317
DB	442	GluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerAsp	461
QY	1318	AACCTTCGAATACGTGCATTCAGACAGACTCAAGGCCACAGCATAGCCAACTTCGCGGT	1377
DB	462	AsnPheGluTyrlleAlaPheLysThrAspSerArgProAsnIleAlaAsnPheAlaGly	481
QY	1378	GAATACTCCGTCTCATATACCTGCGCGAGGAGGTGTCGAAATTCATATGCGCTCCAA	1437
DB	482	GluAsnSerIleIleAspAsnLeuProGluGluValValAlaAsnSerTyrlleGlyLeuPro	501
QY	1438	AGGAGCAGCAGCAGGAGCTTAAAGAACAAACACCCCTTCAAGTTCCTGTTCCACCGTCT	1497
DB	502	ArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSer	521
QY	1498	CAGCATCTCCGAGGCTGTGGCT 1521	
DB	522	GlnGlnSerLeuArgAlaValAla 529	
RESULT 3			
ID	Q647H4	PRELIMINARY;	PRT; 536 AA.
AC	Q647H4;		
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Arachin Any-1.		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	Yan Y., Wang L., Huang S.;		
RT	"cDNA clone of peanut seed storage protein gene."		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY722685; AAU21490.1; -. 30DBF638719AEB78		
SQ	SEQUENCE 536 AA; 61505 MW; 30DBF638719AEB78		
Alignment Scores:			
Pred. No.:	1,02e-163	Length:	536
Score:	2455.00	Matches:	471
Percent Similarity:	93.99%	Conservative:	14
Best Local Similarity:	91.28%	Mismatches:	19
Query Match:	91.23%	Indels:	12
DB:	2	Gaps:	3
US-10-728-323-3 (1-1524) x Q647H4 (1-536)			
QY	1	CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCACGCGCTCAATGCGCAGACCTGAC	60
DB	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnA-gProasp	43
QY	61	AATCGCATGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACAAACCGAGGATTC	120

```

Db      44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAenGlnGluPhe 63
QY      121 GAATGCGCGCGCTCGCTCTCGCTTGGTTCCTCCGCGCAACGCCCTTCGTAGGSCCT 180
Db      64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 83
QY      181 TTCTACTCCAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGGATATCTTTGGGTTG 240
Db      84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY      241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCTCACACAAGGTCTGCTATCTCAG 300
Db      104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
QY      301 TCCCAAGACCAACAGACTCTCCAGGAGAGAGCAAGAGCCAAAGCCAAAGCAGATAGT 360
Db      124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer 143
QY      361 CACCAGAAGGTGCACCGCTTTTCGATGAGGTGATCTCATTTGCAGTTCCCAACCGGTGTGCT 420
Db      144 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY      421 TTCTGGCTCTACAACGACGACGACTGATGTGTTGCTGCTTCTTCTTACTGACACCAAC 480
Db      164 PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY      481 AACAAACGACACCAAGCTTGATCAGTTCCCGCAGAGATTCATTTGGCTGGGAACGGAG 540
Db      184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
QY      541 CAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAIT 600
Db      204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
QY      601 AGCCCATACAGCGCGGAACTAGCTAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db      221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
QY      661 CAGCAGCGCGCAGAGAACGAGCAGCAGCAAGAGAGAAAGCAAGCAAGCTGGAACATCTTC 720
Db      241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGluGlyGlyAsnIlePhe 260
QY      721 AGCGGCTTACGCGCGGAGTTCTCGGAACAAGCTTCAGGTTGACGACAGACAGATAGTG 780
Db      261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu 280
QY      781 CAAAACTTAAGAGCGCAGACCGAGAGTGAAGAAGAGAGGAGCCATTGTGACAGTGGGGA 840
Db      281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
QY      841 GGCCTCAGATCTTGAGCCCGACAGATGAAGAAGAGA-----CGTGCAGCAGAA 885
Db      301 GlyLeuArgIleLeuSerProAspArgPheArgGlnGlnTyrGluArgProAspGlu 320
QY      886 GAAGAGGAATACGATGAAGATGAATATGATACGATCAAGAG-----GATAGA 933
Db      321 GluGluGlnTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnHisAspArg 340
QY      934 AGCGGTGGCAGGGGAGCAGACGCGGGGAATGGTATTGAAGAGAGCATCTGCACCGCA 993
Db      341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
QY      994 AGTGCTAAAAGAACATTGGTAGAAGACAGATCCCTTGACATCTACAAACCTCAAGCTGGT 1053
Db      361 SerPheIysIysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
QY      1054 TCACTCAAACTCCCAACGATCTCAACCTTCTAATCTTAGGTGGCTTGGACCTAGTGCT 1113
Db      381 SerLeuIysThrAlaAsnGluLeuAsnLeuIleLeuArgTrpLeuGlyLeuSerAla 400
QY      1114 GAATATGGAATCTTACAGGAATGCATTTGTCGCTCACTACAAACCAACCGCACAC 1173

```

```

Db      401 GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis 420
QY      1174 AGCATCATATATCGATTGAGGGACGGCTCAGTCAAGTCTGTCAGCAGCAACGGCAAC 1233
Db      421 SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp 440
QY      1234 AGAGTGTACGACGAGGAGCTTCAAGAGGCTCAGCTGCTTGTGTGGTCCACAGAACTTCGCC 1293
Db      441 ArgValPheAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPheAla 460
QY      1294 GTCGCTGGAAAGTCCAGAGCGGAGAACTTCGAATAGCTGCGCATTCAGACAGACTCAAGS 1353
Db      461 ValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAspSerArg 480
QY      1354 CCAGCATAGCAACCTCCGCTGAAACTCCGTCATAGATAACTCCGTCGCGGAGGAGGTG 1413
Db      481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal 500
QY      1414 GTTGCAAAATTCATATGGCTCCAAAGGAGCAGGCAAGCAGCTTAAGAACCAACACCCC 1473
Db      501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsnPro 520
QY      1474 TTCAAAGTCTTCTGTTCCACCGCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db      521 PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla 536

RESULT 4
Q647H3 PRELIMINARY; PRT; 537 AA.
AC Q647H3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arachin Ahy-2.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN (1)_TaxID=3818;
RP SEQUENCE FROM N.A.
RT "cDNA clone of peanut seed storage protein gene."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV722686; AAU21491.1; -.
SQ SEQUENCE 537 AA; 61531 MW; 3BDD034DEA159657 CRC64;

Alignment Scores:
Pred. No.: 1.8e-163 Length: 537
Score: 2451.50 Matches: 474
Percent Similarity: 93.62% Conservative: 10
Best Local Similarity: 91.68% Mismatches: 20
Query Match: 91.10% Indels: 13
Db: 2 Gaps: 4

US-10-728-323-3 (1-1524) x Q647H3 (1-537)
QY 1 CGGCAGCAACCGGAGGAGAACCGCTGCCAGTTCACGCGCTCAATGCGCAGAGACCTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTTGAATACAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
Db 44 AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTGTAGGCT 180
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
QY 181 TTCTACTCCAATGCTCCCGCAGGAGATCTTCATCCAGCAAGGAAGGGGATATCTTTGGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103

```

QY 241 ATATTCCTGGTGTCTCTAGACACTATGAAGAGCTCACACACAGTGTCTCGATCTCAG 300
 Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
 QY 301 TCCCAAGAGACACCAGACGCTCTCAAGGAGAGACCAAGC--CAACAGCAACGAGAT 357
 Db 124 SerGlnArgAlaProArgArgPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 143
 QY 358 AGTCACAGAGGTGACCGTTTCGATGAGGGTGATCTCATTTGCGAGTCTCCACCGGTGTT 417
 Db 144 SerHisGlnLysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 163
 QY 418 GCTTCTCGCTCTACAGACACGACACTGATGTGTTGCTGTTCTTCTTACTGACACC 477
 Db 164 AlaLeuTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 183
 QY 478 AACAAACACCAACACGCTTGATCAGTCTCCAGGAGATTCAATTTGGCTGGGACACG 537
 Db 184 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHis 203
 QY 538 GAGCAGAGATTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAGCAAGATTACCA 597
 Db 204 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 220
 QY 598 TATAGCCCATACGCCCGGAAAGTACGCTAGACAGAGAGCGTGATTTAGCCCTCGA 657
 Db 221 TyrSerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg 240
 QY 658 GGCAGCAGACCGCGCAGACGAGCAGGACAGACAGAAAGCAAGAGGTGGAACATC 717
 Db 241 GlyGlnHisSerArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGln 260
 QY 718 TTCAGCGCTTCACCGCGGAGTCTCTGGAACAGCCTTCAGGTTGACGACAGACAGATA 777
 Db 261 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 280
 QY 778 GTGCAAAACCTAGAGCGGACGAGAGTGTAAGAGAGGAGCGCATTTGACAGTGAGG 837
 Db 281 LeuGlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArg 300
 QY 838 GGAGGCTCAGATCTTTAGCCAGCAGATAGAAAGAGA-----CGTGCGGAC 882
 Db 301 GlyGlyLeuArgIleLeuSerProAspArgGlyArgGlnGlnTyrGluArgProAsp 320
 QY 883 GAAGAAGAGGAATACCATGAAGATGAATATCAATACCATCAAGAG-----GAT 930
 Db 321 GluGluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluArgGlnGlnAsp 340
 QY 931 AGAAGCGTGGCAGGGGAACAGAGCGAGGGGAATGGTATTGAAGAGACGATCTGCACC 990
 Db 341 ArgArgArgGlyArgGlySerArgGlyArgGlyAsnGlyIleGluGluThrIleCysThr 360
 QY 991 GCAAGTCTAAAGAACATTGCTAGAAACAGATCCCTGCATCTACACCCCTCAAGCT 1050
 Db 361 AlaSerValLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAla 380
 QY 1051 GGTTCACCTCAAACTGCCAACGATCTCAACCTCTTAATCTTAGGTGGCTTTGACCTAGT 1110
 Db 381 GlySerLeuLysThrAlaAsnAspLeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSer 400
 QY 1111 GCTGAATATGGAATCTCTACAGAAATGATTTGTTGCTCCTCACTACACACACGCA 1170
 Db 401 AlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAla 420
 QY 1171 CACAGCATCATATATCGATTGAGGGAGCGGCTCAGTGCAGTCTGTGGACAGCAGCGC 1230
 Db 421 HisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGly 440
 QY 1231 AACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTGTGGTCCACAGAACTTC 1290
 Db 441 AsnArgValTyrAspGluGluLeuGlnGlnGlyHisValLeuValValProGlnAsnPhe 460

QY 1291 GCGTCTCGTGGAAAGTCCAGAGCGAGAACTTCGAATAGTGGCATTCAAGACAGACTCA 1350
 Db 461 AlaValAlaGlyLysSerGlnSerAspAsnPheGluTyrValAlaPheLysThrAspSer 480
 QY 1351 AGCCCCAGCATAGCCAACTCGCGGTGAAACTCCGTCATAGATAACCTGCGGAGGAG 1410
 Db 481 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 500
 QY 1411 GTGGTTGCAAAATTCATATGCGCTCCAAAGGAGCAGCAGGAGGCTTTAAGAACAAAC 1470
 Db 501 ValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
 QY 1471 CCCTTCAAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
 Db 521 ProPheLysPheValProProSerGlnGlnSerLeuGlyAlaValAla 537
 RESULT 5
 Q8LKN1 PRELIMINARY; PRT; 538 AA.
 ID Q8LKN1;
 AC Q8LKN1; 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Allergen Arab3/Arab4
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]_TaxID=3818;
 RP SEQUENCE FROM N.A.
 RA Viquez O.M., Konan K.N., Dodo H.W.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF510854; AAM46958.1; -;
 DR HSSP; P04776; 1UCX.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin_region.
 DR InterPro; IPR011051; Rmlc_like_cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
 SQ SEQUENCE 538 AA; 61737 MW; 7AABDD59429709E CRC64;
 Alignment Scores:
 Pred. No.: 6,04e-163 Length: 538
 Score: 2444.00 Matches: 471
 Percent Similarity: 93.63% Conservative: 14
 Best Local Similarity: 90.93% Mismatches: 19
 Query Match: 90.82% Indels: 14
 DB: 2 Gaps: 4
 US-10-728-323-3 (1-1524) x Q8LKN1 (1-538)
 QY 1 CGCGCAACCGGAGGAGAACCGTCCAGCTTCAGCGCTCAATCGCAGACCTGAC 60
 Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
 QY 61 AATCGATTGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
 Db 44 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 63
 QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACCCCTTCGTAGGCCT 180
 Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
 QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGATCTTTGGGTTG 240
 Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
 QY 241 ATATTCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACACAGGTCTGATCTCAG 300

```
Db 104 ILePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
Qy 301 TCCCAAGACACCAAGACGCTCCAGGAGAGACCAAGCCAAAGCAGCAACGAGATAGT 360
Db 124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer 143
Qy 361 CACCAGAAGGTGACACCGTTTCGATGAGGGGTGATCTCATTCAGTCTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
Qy 421 TTCTGGCTCTAACACGACACGACACTGATGTTGTTGCTCTTTCTCTTACTGACACCAAC 480
Db 164 PheTrpMetTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 183
Qy 481 AACCAACGACACCGCTTGATCATGTTCCCCAGGAGATTCAATTTGCTGGACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
Qy 541 CAAGAGTCTTAAAGTACCACGACCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
Qy 601 AGCCCATACGCCCGCAAGCTAGCCTAGACAAGAGAGCGTGAAATTAGCCCTCGAGGA 660
Db 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
Qy 661 CAGCAGCCGCGCAGACGAGCAGCAGACAGACAGAGAAACCAAGCTGGAACATCTTC 720
Db 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnIlePhe 260
Qy 721 AGCGGCTTACCGCGGAGTTCCTGGACAAGCCCTCCAGGTTGACGACACAGATAGTG 780
Db 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspArgGlnIleLeu 280
Qy 781 CAAACCTAAGACGCGAGACCGAGAGTGAAGAGAAGAGGAGCCCATTTGACAGTGGGGA 840
Db 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
Qy 841 GGCTCTAGATCTTGACCCAGATAGAAAGAGA-----CGTGCAGCAGAA 885
Db 301 GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAspGlu 320
Qy 886 GAAGAGGAATACGATGAAGATGAATATGAATACGATCAAG-----GATAGA 933
Db 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluArgGlnGlnAspArg 340
Qy 934 AGCGGTGCGACGGGAAGCAGACGAGCGGGAATGGTATTGAAGAGACGATCTGCACCGCA 993
Db 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
Qy 994 AGTGCTAAAAGAACATTTGTTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGT 1053
Db 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
Qy 1054 TCACTCAAAACTGCCACGAT-----CTCAACCTTCTAATACTTAGTGGCTTGGACCT 1107
Db 381 SerLeuLysThrAlaAsnGluLeuGlnLeuAsnLeuIleLeuArgTrpLeuGlyLeu 400
Qy 1108 AGTGCTGAATATGGAATCTCTACAGGAATGCATTGTTTGTCTCCTCACTACACACCAAC 1167
Db 401 SerAlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsn 420
Qy 1168 GCACACGACATATATATGATTGAGGGAGCGGCTCACGTGCAAGTCTGCGGACAGCAAC 1227
Db 421 AlaHisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsn 440
Qy 1228 GGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACGTGCTGTGGTGCACACAGAAC 1287
Db 441 GlyAspArgValPheAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsn 460
Qy 1288 TTGCGCTGCTGGAAAGTCCCGACGCGAGAACTTGAAATACGTGGCATTCAGACAGAC 1347
```

```
Db 461 PheAlaValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAsp 480
Qy 1348 TCAAGSCCCAGCATAGCCAACTCGCCGCTGAAAACTCCGTCATAGATAACCTGCCGGAG 1407
Db 481 SerArgProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAsnLeuProGlu 500
Qy 1408 GAGGTGGTTGCAAAATTCATATGCGCTCCAAAGGAGCAGCAGGCGAGCTTAAGAAAC 1467
Db 501 GluValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
Qy 1468 AACCCCTTCAAGTCTTCTGTTCCACCGTCTACGAGTCTCCGAGGGCTGTGGCT 1521
Db 521 AsnProPheLysPhePheValProSerGluGlnSerLeuArgAlaValAla 538

RESULT 6
O6T2T4 PRELIMINARY; PRT; 536 AA.
ID O6T2T4;
AC O6T2T4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Storage protein.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Yang H.-X., Wang F., Bi Y.-P.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV439332; AAR02860.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; Rmlc_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
SQ SEQUENCE 536 AA; 61498 MW; 138F3C881BC2C7D5 CRC64;

Alignment Scores:
Pred. No.: 2, 2e-162 Length: 536
Score: 2436.00 Matches: 468
Percent Similarity: 93.60% Conservative: 15
Best Local Similarity: 90.70% Mismatches: 21
Query Match: 90.52% Indels: 12
DB: 2 Gaps: 3

US-10-728-323-3 (1-1524) x O6T2T4 (1-536)
Qy 1 CGGCAGCAACCGGAGAGAACCGGTGCCAGTCTCCAGCGCTCAATCGCGAGACCTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
Qy 61 AATCGCATTTGAATTCAGAGCGCGTTACATTGAGACTTTGGAACCCCAACCAACAGAGATTTC 120
Db 44 AsnArgIleLeuSerGluGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
Qy 121 GAATCGCGCGGCTCGCCCTCTCTCGCTTAGTCTCTCCGCCGCAACGCCCTTCGTAGGCCT 180
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
Qy 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTTC 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgAlaTyrPheGlyLeu 103
Qy 241 ATATTCCCTGGTGTCTAGACACATATGAAGAGCTTCACACACAGGTCTCGATCTCAG 300
Db 104 IlePheLeuGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgHisGln 123
```


Qy	301	TCCCAAGACCA	CAAGACGCTCTCAAGAGAGAACCAAGCCAA	CAGCAACGAGATAGT	360
Db	124	SerGlnArgPro	ProArgPheGlnGlyGlnSerGlnGlnGlnGln	ArgSer	143
Qy	361	CACCAGAAGGT	GCACCGTTTCGATGAGGGTCTCATTCAGTTC	CCACCGGTGGTGC	420
Db	144	HisGlnIysVal	HisArgPheAspGluGlyAspLeuIleAlaValPro	ThrGlyValAla	163
Qy	421	TTCTGGCTCTA	CAACGACCAACGACACATCATGTGCTGTGCT	TTCTTACTGACACCAAC	480
Db	164	PheTrpMet	TyrAsnAspHisAspThrAspValValAlaValSer	LeuThrAspThrAsn	183
Qy	481	AACAACGACAA	CCAGCTTGATCATGTTCCCGAGGAGATTCAT	TTGGTGGTGGGAACACGGAG	540
Db	184	AsnAsnAsp	AsnGlnLeuAspGlnPheProArgArgPheAsnLeu	AlaGlyAsnHisGlu	203
Qy	541	CAAGAGTTCTT	TAAGGTACCAGCAACAAAGCAGACAGACAA	AGACGACCAAGAGCTTACCATAT	600
Db	204	GlnGluPhe	LeuArgTyrGlnGln-----GlnSerArgArg	ArgSerLeuProTyr	220
Qy	601	AGCCCATACAG	CCCGCAAGATCAGCCTAGACAAGAGAGCGT	GAATTTAGCCCTCGAGGA	660
Db	221	SerProTyr	SerProGlnThrGlnProIysGlnGluAspArg	GluPheSerProArgGly	240
Qy	661	CAGCACAGCCG	CAGAGAACGACGACAGGACAGAGAAAGAA	CAAGAGTGGAAACATCTTC	720
Db	241	GlnHisGly	ArgArgGluArgAlaGlyGlnGlnGlnGlnGln	AsnGluGlyGlyAsnIlePhe	260
Qy	721	AGCGGCTTCA	CGCGGAGTTCCTGGAAACAGCCTTCAGGTT	GACGACGACAGATAGTG	780
Db	261	SerGlyPhe	ThrProGluPheLeuAlaGlnAlaPheGlnVal	AspAspArgGlnIleLeu	280
Qy	781	CAAACTTAAG	AGGCGGACGACCGAGGTGAAGAGAGGAGGC	ATGTGTGACAGTGGAGGGA	840
Db	281	GlnAsnLeu	argGlyGluAsnGluSerAspGluGlnGlyAla	IleValThrValArgGly	300
Qy	841	GGCCTCAGA	ATCTTTGAGCCCAAGATAGAAAGAGA-----	CGTCCGACGAA	885
Db	301	GlyLeuArg	IleLeuSerProAspArgGlyIysArgGlnGln	TyrGluArgProAspGlu	320
Qy	886	GAAGAGAA	TACGATGAGATATGATATACCATCAAGAG-----	GATAGA	933
Db	321	GluGluGlu	TyrAspGluAspGluTyrGluTyrAspGluGlu	GluArgGlnGlnAspArg	340
Qy	934	AGGCGTGGC	AGGGAAGCAGAGGCGAGGGGGAATCGTATT	GAAGAGACGATCGCACGCA	993
Db	341	ArgArgGly	argGlySerArgGlySerGlyAsnGlyIleGlu	GluThrIleCysThrAla	360
Qy	994	AGTGCTAAA	AGAAATGTTGTPAGAAACAGATCCCTGACAT	CTACACCCCTCAAGCTGGT	1053
Db	361	SerPheIys	IysAsnIleGlyArgAsnArgSerProAspIle	TyrAsnProGlnAlaGly	380
Qy	1054	TCACTCAAA	CTGCAACGATCTCAACCTTCTATATCTTAGT	TGGCTTGGACCTAGTGCT	1113
Db	381	SerLeuIys	ThrAlaAsnGluLeuAsnLeuIleLeuArgTrp	LeuGlyLeuSerAla	400
Qy	1114	GAATATG	GAATCTCTCAGGAATGCATTTGTTGCTCTAC	TACTACACCAACGACGACAC	1173
Db	401	GluTyrGly	AsnLeuTyrArgAsnAlaIleuPheValPro	HisTyrAsnThrAsnAlaHis	420
Qy	1174	AGCATCATAT	TCGATTCAGGGGACGGGCTCACGTGCAAGT	CGTGGACAGCAACCGCAAC	1233
Db	421	SerIleIle	IleTyrAlaLeuArgGlyArgAlaHisValGln	ValValAspSerAsnGlyAsp	440
Qy	1234	AGATGTAC	AGNAGGAGCTTCAAGAGGGTCACTGCTTGT	TGGTGCCACAGAACTTCGCC	1293
Db	441	ArgValPhe	AspGluGluLeuGlnGlnGlyHisValLeu	ValValProGlnAsnPheAla	460
Qy	1294	GTCGCTG	GAAGTCCCAGAGCGGAACTTCGAATACGT	TGGCATTCACAGACAGACTCAAGG	1353
Db	461	ValAlaGly	IysSerGlnSerGlnAsnPheGluTyrVal	IleAlaPheIysThrAspSerArg	480
Qy	1354	CCAGCATAG	CCAACTCCGCGGTGAAAACCTCCGCTCAT	AGATATACTCCGCGGAGAGGTG	1413

Db 481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal 500
 Qy 1414 GTTGCAAAATCATATGGCCCTCCAAAGGGAGCAGGAGCGCTTAAGAAACAACACCCC 1473
 Db 501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuIleuysAsnAsnPro 520
 Qy 1474 TTCAAGTCTTCGTTCCACCGTCTCAGCAGCTCCGAGGGCGTGTGGCT 1521
 Db 521 PheIysPhePheValProProSerGluGlnSerLeuArgAlaValAla 536
 RESULT 7
 Q9SQH7 PRELIMINARY; PRT; 530 AA.
 ID AC Q9SQH7
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Glycinin.
 GN Names:Ararh4;
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 ON NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Virginia;
 RX MEDLINE=9406463; PubMed=10474031; DOI=10.1159/000024203;
 RA Kieber-Janke T., Crameri R., Appenzeller U., Schlaak M., Becker W.M.;
 RT "Selective cloning of peanut allergens, including profilin and 2S
 RT albumins, by phage display technology.";
 RL Int. Arch. Allergy Immunol. 119:265-274(1999).
 DR EMBL; AF086821; AAD47382.1; -;
 DR HSP; P04776; IUCK.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR011051; RmlC like cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00150; Cupin; 2.
 DR PRINTS; PR00439; IISGLOBULIN.
 SQ SEQUENCE 530 AA; 61011 MW; 864E349A31642A32 CRC64;
 Alignment Scores:
 Pred. No.: 2,48e-161 Length: 530
 Score: 2421.00 Matches: 469
 Percent Similarity: 94.09% Conservative: 9
 Best Local Similarity: 92.32% Mismatches: 29
 Query Match: 89.97% Indels: 2
 DB: Gaps: 0
 US-10-728-323-3 (1-1524) x Q9SQH7 (1-530)
 Qy 1 CGGCAGCAACCGGAGGAGAACGGCTGCCAGTCTCCAGCGCTCAATGCCAGACCTGAC 60
 Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
 Qy 61 AATCGCATTAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
 Db 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
 Qy 121 GAATGGCGGGCTGCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGSCCT 180
 Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 83
 Qy 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGTTG 240
 Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
 Qy 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGACGCTTCACACAAGGTCTGCTCCTCAG 300

Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgTyrGln 123
QY 301 TCCCAAGACACCAAGACCTCCAGGAGAGAGACCAAGCCAAAGCCAAAGCAGACAGATAGT 360
Db 124 SerGlnArgProProArgArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 143
QY 361 CACCAGAGGTGCACCGTTCGATAGGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAsnGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGGCTCTACAGACACGACGACCTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 164 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACACGACACACCGCTTGATCAGTTCCCCAGGAGATTCATTTGGCTGGGAACCGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223
QY 601 AGCCCATACAGCCGCGMAAGTACGCTAGACAGAGAGCGTGAATTTAGCCCTCCAGGA 660
Db 224 SerProTyrSerProHisSerArgProArgGluGluArgGluPheArgProArgGly 243
QY 661 CAGCACAGCCGACAGACAGCAGCAGACAGACAGAAAGAAACGAAGTGGAAACATCTTC 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAspGluGlyAsnIlePhe 263
QY 721 AGCGGCTTACCGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGACAGACAGATAGTG 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAAACCTAGACGCGACCGAGAGTAGAGTAGAGAGAGAGCCATTGTGACAGTGAGGGA 840
Db 284 GlnAsnLeuTrpGlyGluAsnGlnSerGluGlnGluGlyAlaIleValThrValArgGly 303
QY 841 GGCTCAGAACTTTGAGCCAGATAGAAAGAGACGTCGCCAGCAAGAGGAATACGAT 900
Db 304 GlyLeuArgIleLeuSerProAspGlyThrArgGlyAlaAspGluGluGluTyrAsp 323
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAGGCGTGGCAGGGGAAGCAGAGCAGG 960
Db 324 GluAspGlnTyrGluTyrHisGluGlnAspGlyArgArgGlyArgGlySerArgGlyGly 343
QY 961 GCGAATGGTATTGAGAGACGATCGCACCGCAGTGCTTAAAGAACATTTGTTAGAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaCysValLysLysAsnIleGlyGlyAsn 363
QY 1021 AGATCCCTGACATCTACAAACCTCAAGC-TGGTTCACTCAAACTGCCAACGATCTCAA 1079
Db 364 ArgSerProHisIleTyrAspProGlnArgTrpPheThrGlnAsnCys-HisAspLeuAs 383
QY 1080 CCTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGC 1139
Db 383 nLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAl 403
QY 1140 ATTGTTTGTCTGCTACTACAAACCAACGACACAGCATCATATATCGATTGAGGGGACG 1199
Db 403 aLeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyAr 423
QY 1200 GGCTCAGCTCAAGTCTGACAGACAGCGCAACAGAGTGTACGACGAGGAGCTTCAAGA 1259
Db 423 gAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluLeuGlnG 443
QY 1260 GGTCACGCTCTGTGGTCCACAGAACTTCGCGCTCGTGGAAAGTCCAGAGCGAGAA 1319
Db 443 uGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAs 463
QY 1320 CTTTCGATAGCTGGCATTCAGACAGACTCAAGCCCGCATAGCCAACTCCCGCGTGA 1379
Db 463 nPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnPheAlaGlyG 483

QY 1380 AAATCCGTCATAGATAACCTCCGAGGAGGTGGTTGCAAAATTCATATGGCTCCAAAG 1439
Db 483 uAsnSerPheIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuProAr 503
QY 1440 GGAGCAGGCAAGCGAGCTTAAAGAACAAACCCCTTCAAGTTCCTCGTTCCACCGTCTCA 1499
Db 503 gGluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProPropheG 523
QY 1500 GCAGTCTCCGAGGGCTGTGGCT 1521
Db 523 nGlnSerProArgAlaValAla 530
RESULT 8
Q61WG5 PRELIMINARY; PRT; 510 AA.
AC Q61WG5; Q61WG5; (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycinin (Fragment).
OS Arachis hypogaea (peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Kang I.-H., Gallo-Meagher M.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY618460; AAT39430.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SBED_STORAGE; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 510 AA; 58061 MW; F6569EF13D1A12BD CRC64;
Alignment Scores:
Pred. No.: 3,77e-120 Length: 510
Score: 1834.00 Matches: 364
Percent Similarity: 78.57% Conservative: 43
Best Local Similarity: 70.27% Mismatches: 71
Query Match: 68.15% Indels: 40
DB: 2 Gaps: 8
US-10-728-323-3 (1-1524) x Q61WG5 (1-510)
QY 1 CGGCACCAACCGAGGAGAACCGTGCAGTTCACGCGCTCAATGCGCAGAGACCTGAC 60
Db 22 ArgGlnGlyGlyGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 41
QY 61 AATCGATTGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACCAACAGAGATTC 120
Db 42 AsnArgIleLeuSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 61
QY 121 GAATCGCGCGGCTCGCCCTCTCTCGTCTAGTCTCCGCGCAAGCCCTTCGTAGGCCT 180
Db 62 GlnCysAlaGlyValAlaLeuSerArgThrValLeuArgAsnAlaLeuArgArgPro 81
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCGACGAGGAGGATCTTTGGGTG 240
Db 82 PheTyrSerAsnAlaProLeuGluIleTyrValGlnGlnGlySerGlyTyrPheGlyLeu 101
QY 241 ATATTCCCTGGTGTCTCTAGACACTATGAGAGCCTCACACAGAGTTCGTGATCTCAG 300
Db 102 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGluGlyArgTyrGln 121
QY 301 TCCCAAGACCAACCAAGACGTCTCCAA-----GGGAAGACCAAGCCCAACGACGA 354

```
122 SerGlnYsProSerArgArgPheGlnValGlyGlnAspPProSerGlnGlnGln 141
135 GATAGTCACCAAGAGTGCACCGTTTCGATGAGGGTGAATCTCATTCAGTTCACCGGT 414
142 AspSerHisGlnYsValHisArgPheAspGluGlyAspLeuIleAlaValProThrGly 161
145 GTTGTCTTCTGCTCTACACGACACGACGACGACGACGACGACGACGACGACGACGAC 474
162 ValAlaPheTrpMetTyrAsnAspGluAspValValThrValThrLeuSerAsp 181
175 ACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 534
182 ThrSerSerIleHisAsnGlnLeuAspGlnPheProArgArgPheTyrLeuAlaGlyAsn 201
195 ACGAGCAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACAAAGCTTA 594
202 GlnGluGlnGluPheLeuArgTyrGlnGlnGlnGlySer----- 215
215 CCATATAGCCATACACCCCGCAAGTCAGCTTCCAGGAGATTCATTTGGCTGGGAAAC 634
216 -----ArgProHisTyrArgGlnIleSerPro 224
225 CGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 714
225 ArgValArg-----GlyAspGluGlnGluAsnGluGlySerAsn 237
237 ATCTTACGGCTTCACCGCGAGTTCCTGGAACAAAGCCTTCAGAGTTGACGACAGACAG 774
238 IlePheSerGlyPheAlaGlnGluPheLeuGlnHisAlaPheGlnVal---AspArgGln 256
245 ATAGTCAAAACCTAGAGCGGAGCAGCAGAGTGAAGAGAGGAGGAGCATTGTGACAGTG 834
257 ThrValGluAsnLeuArgGlyGluAsnGluArgGluGlnGlyAlaIleValThrVal 276
285 AGGGAGCGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTGC----- 879
277 LysGlyGlyLeuArgIleLeuSerProAspGluAspGluSerSerArgSerProPro 296
280 GACGAAGAAGAGGAATACGATGAAGATGAATATATACGATGAAGAGGATAGAGCGGT 939
297 SerArgArgGluGluPheAspGluAspArg-----SerArgProGlnGlnArg 312
300 GGCAGG-----GGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 990
313 GlyLysTyrAspGluAsnArgArgGlyTyrLysAsnGlyIleGluGluThrIleCysSer 332
325 GCAAGTCTAAAGAAACATTGGTAGAAACAGATCCCTCAGATCTACAAACCTCAAGCT 1050
333 AlaSerValLysLysAsnLeuGlyArgSerSerAsnProAspIleTyrAsnProGlnAla 352
351 GGTCTACTCAAACTGCAACGATCTCAACCTTCTAATCTAGGTGGCTTGGACCTTAGT 1110
353 GlySerLeuArgSerValAsnGluLeuAspLeuProIleLeuGlyTrpLeuGlyLeuSer 372
381 GCTGAATATGAATCTCTACAGGATGATTTGTCTGCTCCTCCTACACACCAAGCA 1170
373 AlaGlnHisGlyThrIleTyrArgAsnAlaMetPheValProHisTyrThrLeuAsnAla 392
395 CACAGCATATATATGATTGAGGGAGCGGCTCAGCTGCAAGTCTGGGACAGCAACGGC 1230
395 HisThrIleValValAlaLeuAsnGlyArgAlaHisValGlnValValAspSerAsnGly 412
421 AACAGAGTGCACGAGGAGCTTCAAGAGGGTCAGTGTCTGTGGTCCACAGAACTTC 1290
413 AsnArgValTyrAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPhe 432
429 GCGTGTCTGGAAGTCCACGAGGAGAACTTCGATACGTGGCTTCAAGACAGACTCA 1350
433 AlaValAlaLysAlaGlnSerGluAsnTyrGluTyrLeuAlaPheLysThrAspSer 452
435 AGCCGAGCATAGCAACCTCGCGGTGAACCTCGCTCATAGTAACCTCGCGGAGGAG 1410
```

```
453 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 472
1411 GTGGTTGCAAAATTCATATATGCTTCCAAAGGAGCAGCAGGAGGAGGAGGAGGAGGAG 1470
473 ValValAlaAsnSerTyrArgLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 492
1471 CCCTTCAAGTTCTTCGTTCCACCG---TCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
493 ProPheLysPheValProPheAspHisGlnSerMetArgGluValAla 510

RESULT 9
Q647H2 PRELIMINARY; PRT; 484 AA.
ID Q647H2 Q647H2
AC Q647H2 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arachin Ahv-3.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV722687; AAU21492.1; -.
SQ SEQUENCE 484 AA; 54568 MW; 5A3E950752E89D2D CRC64;

Alignment Scores:
Pred. No.: 2,58e-107 Length: 484
Score: 1651.00 Matches: 337
Percent Similarity: 75.25% Conservative: 46
Best Local Similarity: 66.21% Mismatches: 76
Query Match: 61.35% Indels: 50
DB: 2 Gaps: 8

US-10-728-323-3 (1-1524) x Q647H2 (1-484)
QY 1 CGGCAGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 24 ArgGlnGlnGlyGluGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 43
QY 61 AATCGCATTTGAATTCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACCAACCAAGAGTTC 120
DB 44 AsnCysIleGluSerGluGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
DB 64 GlnCysAlaGlyValAlaLeuSerArgPheValLeuArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleTyrGlnGlySerGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCCTCACACCAAGGTCGTCTGATCTCAG 300
DB 104 IlePheProGlyCysProGlyThrPheGluGluPro---IleGlnGly---SerGluGln 121
QY 301 TCCCAAGACCAACAGACGCTTCCAAAGGAGAGCAACCAAGCAACCAAGCAACGAGATAGT 360
DB 122 PheGlnArgProSerArgHisPheGlnGlyGlnAspGlnSerGlnArgProLeuAspThr 141
QY 361 CACCAGAGGTCGACCGTTTCGATGAGGCTGATCTCATTCGAGTTCACCGCGTGTGCT 420
DB 142 HisGlnLysValHisGlyPheArgGluGlyAspLeuIleAlaValProHisGlyValAla 161
QY 421 TTCTGCTCTACACGACCAACGACGACGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 162 PheTrpIleTyrAsnAspGlnAspThrAspValValAlaIleSerValLeuHisThrAsn 181
```

Qy	481	AACAAACGACAAACCAAGCTTGTATCAGTTCCCCAGGAGATTCAATTTGGCTGGGAACACCGAG	540
Db	182	SerLeuHisAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyLysGlnGlu	201
Qy	541	CAAGAGTTCTTAAGGTACCAGCAACAAGC---AGACAAAGACACAGNAGNAGCTTACCA	597
Db	202	GlnGluPheLeuA-rGlyrGlnGlnArgSerGlyrGlnSer	215
Qy	598	TATAGCCCATACAGCCCGCAAGTCAGCTAGACAAGAAGACGCTGAATTTTAGCCCTCGA	657
Db	216	-----ProlysGlyGluGlnGlu-----	222
Qy	658	GGACAGCACACCGCAGAGAAACGACGAGGACAAGAAGAAAAACGAAGTGTGAAAACATC	717
Db	223	-----GlnGluGlnGluAsnGluGlyGlyAsnVal	232
Qy	718	TTACAGCGGCTTACGCCGAGTTCCTCGAACACAGCCTTCAGGTTTGACGACAGACAGATA	777
Db	233	PheSerGlyPheSerThrGluPheLeuSerHisGlyPheGlnValAsnGlu---AspIle	251
Qy	778	GTGCAAAACCTTAAGAGCGACACGACGAGTGTCAAGAGAGGAGGAGCCATTGTGCACAGTCAGG	837
Db	252	ValArgAsnLeuA-rGlyrGluAsnGluA-rGluGlnGlnGlyAlaIleValThrValLys	271
Qy	838	GGAGGCTCAGAATCTTGAGCCCAAGATAGAAAGAGACGTGCCGACGAAGAAGAGGAATAC	897
Db	272	GlyGlyLeuSerIleLeuValProPro-----	280
Qy	898	GATCAAGATGAATATGAATACGATAGAGGATAGAAGCGCTGGCAGGGAAGCAGAGGC	957
Db	281	-----GluTrpArgIleSerThrGlnGlnProGlyA-rGlyAspLysAsp	295
Qy	958	AGGGGGAATGTATTGAAGAGACGATCTGCACCGCAAGTGTCAAAAAGAACATTGGTAGA	1017
Db	296	PheAsnAsnGlyIleGluGluThrIleCysThrAlaThrValLysMetAsnIleGlyLys	315
Qy	1018	AACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAAACCTGCCAAGCATCTC	1077
Db	316	SerThrSerAlaAspIleTyrAsnProGlnAlaGlySerValA-rGlyrThrValAsnGluLeu	335
Qy	1078	AACCTTCTAATACTTAGTGCTTGGACCTAGTCTGCAATGATGAAATCTCTACAGGAAT	1137
Db	336	AspLeuProIleLeuAsnA-rGluGlyLeuSerAlaGluTyrGlySerIleHisA-rGasp	355
Qy	1138	GCATTGTTGTGCTCACTACAACCAACGACACACGATCATATATCATGATTGAGGGA	1197
Db	356	AlaMetPheValProHisTyrAsnMetAsnAlaAsnSerMetIleTyrAlaLeuHisGly	375
Qy	1198	CGGCTCACGTGCAAGTCTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAA	1257
Db	376	GlyAlaHisValGlnValValAspCysAsnGlyAsnA-rGValPheAspGluGluLeuGln	395
Qy	1258	GAGGCTCACGTCCTTGTGGTCCACAGAACTTCGCCGCTCGTGGAAGTCCAGAGCGAG	1317
Db	396	GluGlyGlnSerLeuValValProGlnAsnPheAlaValAlaLysSerGlnSerGlu	415
Qy	1318	AACTTTGAATACGTGGCATTCAGACAGACTCAAGGCCGACATCATATATCGATCGCCGT	1377
Db	416	HisPheLeuTyrValAlaPheLysThrAsnSerArgAlaSerIleSerAsnLeuAlaGly	435
Qy	1378	GAAACCTCCGTATAGATTAACCTGCCGAGGAGGTGCTGCAATTCATATGGCTCCAA	1437
Db	436	LysAsnSerTyrMetTrpAsnLeuProGluAspValValAlaAsnSerTyrGlyLeuGln	455
Qy	1438	AGGAGAGGCAAGCGCAGCTTAAGAAACAACACCCCTTCAAGTTCTCTCGTTCACCGC	1494
Db	456	TyrGluGlnAlaA-rGlnLeuLysAsnAsnAsnProPheThrPheLeuValProGln	475
Qy	1495	TCTCAGCAGTCTCCAGGCTGTGGCT	1521
Db	476	AspSerGlnMetIleA-rGThrValAla	484

```

RESULT 10
GLC3 SOYBN
ID - GLC3 SOYBN STANDARD; PRT; 481 AA.
AC P11828;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycinin G3 precursor [Contains: Glycinin A subunit; Glycinin B subunit].
DN Name=GY3;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; OC Eustersids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. NCBI_TaxId=3847;
OX [1]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296500; PubMed=2740231;
RA Cho T.-J., Nielsen N.C.;
RT "The glycinin G3 gene from soybean.";
RL Nucleic Acids Res. 17:4388-4388(1989).
[2]
RN DISCUSSION OF SEQUENCE.
RX MEDLINE=92393191; PubMed=2485233;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J., Fischer R.L., Sims T.L., Drees G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
CC -1- SUBUNIT: Glycinin is the major seed storage protein of soybean. CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.
CC -1- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
CC EMBL; X15123; CAA33217.1; -.
DR PIR; S04605; S04605.
DR HSPG; P04776; IPIXZ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; Rmlc like cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 296 Glycinin A subunit.
FT CHAIN 297 476 Glycinin B subunit.
FT PROPEP 477 481
FT DISULFID 107 303 Interchain (By similarity).
SQ SEQUENCE 481 AA; 54241 MW; 5F3C3148DF6241A7 CRC64;
-----
Alignment Scores:
Pred. No.: 7,33e-97 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 1 Gaps: 7
US-10-728-323-3 (1-1524) x GLC3 SOYBN (1-481)

```


Db 238 ArgAenLeuGlnGlyGluAenGluGluAspSerGlyAlaIleValThrValLysGly 257
 Qy 841 GGCTCAGAACTTTGACGCCAGATAGAAAGAGAGCTGCCGACGAAAGAGGAATACGAT 900
 Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluGluAspAspAsp 277
 Qy 901 GAAGATGAA-----TATGAATACGATCAAGAGGATAGAAAGCGTGGCAGGGGA 948
 Db 278 GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
 Qy 949 AGCAGAGCGAGGGGGAATGGTATTCAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAC 1008
 Db 298 SerArg-----AenGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
 Qy 1009 ATTGGTGAACAGATCCCTGACATCATCAACCTCAAGCTGGTTCACCTCAAACTGCC 1068
 Db 315 IleGlyGlnAenSerProAspIleThrAenProGlnAlaGlySerIleThrAla 334
 Qy 1069 AACGATCTCAACCTTCTAATCTTAGTGGCTTGGACCTAGTGTGTAATGGAAATCTC 1128
 Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTrpGlySerLeu 354
 Qy 1129 TACAGGAATGATTTGTTGCTGCTCACTACCAACCAAGCAGACATCATATATCGA 1188
 Db 355 ArgLysAenAlaMetPheValProHisThrLeuAenAlaAenSerIleIleThrAla 374
 Qy 1189 TTGAGGGGACGGGCTCAGTGCAGTCAAGTGTGACAGCAACGGCAACAGAGTGTACGACGAG 1248
 Db 375 LeuAenGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGly 394
 Qy 1249 GAGCTTCAAGAGGCTCAGCTGCTGTGTCGCCAGCAACTCGCGCTCGTGGAAAGTCC 1308
 Db 395 GluLeuGlnGlyGlyValLeuIleValProGlnAenPheAlaValAlaAlaLysSer 414
 Qy 1309 CAGAGCGAGAACTTCGAATACGTGGCTTCAAGACAGACTCAAGCGCCAGCATAGCCAAAC 1368
 Db 415 GlnSerAenAenPheGluTrpValSerPheLysThrAenAspArgProSerIleGlyAen 434
 Qy 1369 CTCGCCGGTGAACACTCCGTCAATAGATAAATCTGCCGAGGAGGTGGTGCAAATTCATAT 1428
 Db 435 LeuAlaGlyAlaAenSerLeuLeuAenAlaLeuProGluGluValIleGlnHisThrPhe 454
 Qy 1429 GGCTTCAAGGGAGCAGGCAAGCAGCTTAAAGACAAACCCCTTCAAGTCTTCTCGTT 1488
 Db 455 AenLeuLysSerGlnGlnAlaArgGlnValLysAenAenAsnProPheSerPheLeuVal 474
 Qy 1489 CCACCGCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
 Db 475 ProGlnGluSerGlnArgArgAlaValAla 485

RESULT 12
 ID -GLC1 SOYBN STANDARD; PRT; 495 AA.
 AC P04776;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Glycinin G1 precursor [Contains: Glycinin A1A subunit; Glycinin BX subunit].
 GN Name=GV1;
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Dare; TISSUE=Leaf;
 RA MEDLINE=92393391; PubMed=2485233;
 RX Nielsen N.C., Dickson C.D., Cho T.J., Thanh V.H., Scallan B.J.,
 RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
 RT "Characterization of the glycinin gene family in soybean.";
 RL Plant Cell 1:313-328(1989).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Dare; TISSUE=Leaf;
 RA MEDLINE=89296498; PubMed=2740229;
 RX Sims T.L., Goldberg R.B.;
 RT "The glycinin Gyl gene from soybean.";
 RL Nucleic Acids Res. 17:4386-4386(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bonminori;
 RX MEDLINE=86041867; PubMed=2997720;
 RA Negoro T., Momma T., Fukazawa C.;
 RT "A cDNA clone encoding a glycinin A1A subunit precursor of soybean.";
 RL Nucleic Acids Res. 13:6719-6731(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Utsunomi S., Kohno M., Mori T., Kito M.;
 RT "An alternate cDNA encoding glycinin A1A Bx subunit.";
 RL J. Agric. Food Chem. 35:210-214(1987).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Urade R., Nakatani H., Nakano C.;
 RT "mRNA of soybean proglycinin A1A1b subunit.";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -i- FUNCTION: Glycinin is the major seed storage protein of soybean.
 CC -i- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.
 CC -i- PTM: The precursor is post-translational processed to form a covalently linked A1A-BX subunit complex.
 CC -i- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M36686; AAA33966.1; -;
 DR EMBL; X15121; CAA33215.1; -;
 DR EMBL; X02985; CAA26723.1; -;
 DR EMBL; AB113349; BAC78522.1; -;
 DR PIR; A23497; FWSYG2.
 DR PIR; S10851; S10851.
 DR PDB; 1EXZ; X-ray; A/B/C=20-495.
 DR PDB; 1UCX; X-ray; A/B/C=20-495.
 DR PDB; 1UDI; X-ray; A/B/C=20-495.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR011051; RmlC like cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
 KW 3D-structure; Multigene family; Seed storage protein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 306 Glycinin A1A subunit.
 FT PROPEP 307 310
 FT CHAIN 311 490 Glycinin BX subunit.
 FT PROPEP 491 495
 FT DISULFID 107 317 Interchain (By similarity).
 FT CONFLICT 42 42 D -> G (in Ref. 3).
 FT CONFLICT 108 108 P -> S (in Ref. 3).
 FT CONFLICT 136 136 F -> S (in Ref. 3).
 FT CONFLICT 360 360 E -> G (in Ref. 3).
 FT TURN 30 31
 FT STRAND 39 39
 FT STRAND 43 47
 FT TURN 48 49
 FT TURN 50 54

QY 1 CGGCAGCAACCGGAGGAGAACCGGTGCCAGTTCACGCGCCTCAATGGCAGAGACCTGAC 60

Db	23	ArgGluGlnProGlnGlnAenGluCysGlnIleGlnArgLeuAenAlaLeuLysProGly	42
Qy	61	AATCGCATTGAATCAGAGCGGTTACATTAGACATTGGAAACCCCAACACACGAGGATTC	120
Db	43	AenArgIleGluSerGluGlyGlyPheIleGluThrTrpAenProAenAenLysProPhe	62
Qy	121	GAATGGCGCGGCTCGCCCTCTCCGTTAGTCTCCGCCGCAACGCCCTTCGTAGGCGCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAenArgAenAlaLeuA-ArgPro	82
Qy	181	TTCTACTCCAAATGTCCTCCAGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTG	240
Db	83	SerTyrThrAenAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
Qy	241	ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCCTCACACAAAGTCTGTCGATCTCAG	300
Db	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	119
Qy	301	TCCCAAGACCACCAAGACGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT	360
Db	120	SerSerArgPro-----GlnAenArg	126
Qy	361	CACCAGAAGGTGCACCGCTTTCGATGAGGGTGATCTCATTCAGTTCACCAGGTGTGTCT	420
Db	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
Qy	421	TTCTGGCTCTACAAACGACACGACACTGATGTTGTTGCTGCTTTCTCTTACTGACACCAAC	480
Db	147	TyrTrpMetTyrAenAenGluAenThrLeuValValAlaValSerLeuIleAenThrAen	166
Qy	481	AACAAACGACACGACTTGATCAGTTCCTCCAGAGATTCAATTTGGCTGGGAACACGGAG	540
Db	167	SerPheGlnAenGlnLeuAenGlnMetProArgPheTyrLeuAlaGlyAenGlnGlu	186
Qy	541	CAAGAGTCTTAAGGTACACGACGACAAAGCAGACAGACAGACGAAGAAGCTTACCATAT	600
Db	187	GlnGlnPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----	199
Qy	601	AGCCCATACAGCCGCAAGTCAAGCTAGACAAGAAGAGCGTGAAATTTAGCCCTCCGAGGA	660
Db	200	-----Gly	200
Qy	661	CAGCACAGCCGACAGAAACGAGCAGGACGACAGAAAGAAACGAAAGTGGAAACATCTTC	720
Db	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAenGluGlyGlySerMetLeu	220
Qy	721	AGCGGCTTACGCGGAGTCTCTGGAAACAGCCCTCCAGCTTCACGACACACAGATAGTG	780
Db	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
Qy	781	CAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGGAGCCATTGTGACAGTGGAGGGA	840
Db	240	ArgLysLeuGlnGlyGluAenGluGluGluGlyAlaIleValThrValLysGly	259
Qy	841	GGCCTCAGAATCTTGAGCCCA-----GATAGAAAGAGACGTGCCGACGAGAAAGAG	891
Db	260	GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu	279
Qy	892	GAATACGATCAAGATGNAATGAATACGATGAGAGGATAGAGGCGTGCACGGGGNAGC	951
Db	280	LysProAen-----CysAenGluLysAenLysHisCysGlnSerGlnSer	294
Qy	952	AGAGGACGGGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAACATT	1011
Db	295	Arg-----AenGlyIleAspGluThrIleCysThrMetArgLeuArgHisAlaIle	311
Qy	1012	GGTAGAAACAGATCCCTCGACATCTACAACCCCTCAAGCTGGTTCTCACTCAAACTGCCAAC	1071
Db	312	GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrThrAlaThr	331
Qy	1072	GATCTCAACCTTCTATACTTAGGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTAC	1131
Db	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351

```

QY 1132 AGGAATGCATTTGTTGCTCACTACAACCAACGACACACAGCATCATATATCGATTG 1191
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 LysAsnAlaMet PheValProHisTyrAsnLeuAsnAlaAsnSerIleTyrAlaLeu 371
QY 1192 AGGGAGCGGCTCAGCTGCAGTCGTGGAGACAGCAACAGAGTGTACGACGAGGAG 1251
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGTCACGTGCTGGTGGCCACACAACTTCGCGTCGCTGGAAAGTCCAG 1311
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCGAGAATCTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTTC 1371
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCCGGTGAACACTCCGTCATAGATACCTGCGGAGAGGTGGTGCATATCATATGCG 1431
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 AlaGlyAlaAsnSerLeuLeuAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCCAAGGAGCAGCAGCAGGAGCTTAAGAACAAACCCCTTCAAGTTCCTCGTTCCA 1491
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
472 ProLysGluSerGlnArgValValAla 481

RESULT 14
Q852U4 PRELIMINARY; PRT; 482 AA.
AC Q852U4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycinin A1bB2-784.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Matsuura; TISSUE=Seed;
RA Fukazawa C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030495; BAC55938.1; -.
DR HSP; P04776; 1FXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN 1.
SQ SEQUENCE 482 AA; 54298 Mw; 79086863D946BC0D CRC64;

Alignment Scores:
Pred. No.: 6,49e-96 Length: 482
Score: 1488.50 Matches: 292
Percent Similarity: 71.37% Conservative: 72
Best Local Similarity: 57.25% Mismatches: 93
Query Match: 55.31% Indels: 53
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x Q852U4 (1-482)

QY 1 CGGCACAACCGGAGGAGAACGGTCCAGTTCACGCGCTCAATCGCGAGAGACCTGAC 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProGly 42

```

```

QY 61 AATCCGATTGAATCAGAGGCGGTTCATTGAGACTTGGAAACCCCAACACGAGGATTC 120
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCT 180
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 GlnCysAlaGlyValAlaLeuSerArgTyrThrLeuIleArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGATACATTTGGGTG 240
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 SerTyrThrAsnAlaProGlnGlnIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTGTCTAGACATATGAAGAGCTTCACACAAGTTCGTGCATCTCAG 300
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACACCAAGACGCTCCAGGAGAACCAAGCAACAGCAACGACGATAGT 360
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyLeuAla 146
QY 421 TTCTGCTCTACAACGACCAACGACACTGATGTGTGTGTCTTCTTACTGACACCAAC 480
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 TyrTrpMetTyrAsnAsnGluAspThrLeuValValAlaValSerIleIleAspThrAsn 166
QY 481 AACACGACCAACGAGTTCATCAGTTCCTCCAGGAGATTCATTTGGTGGGAACACGGAG 540
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGln 186
QY 541 CAAGAGTCTTAAGGTACCAGCAACAAGACAGCAAGAGCAAGCAAGCAAGCAAGTTCACCAT 600
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 GlnGluPheLeuGlnTyrGlnSerGlnLysGlnGlnGly-----Gly 200
QY 601 AGCCCATACAGCCCGCAAGTCAAGTACAGCAAGAGAGCGTGAATTTAGCCCTCGAGA 660
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 -----Gly 200
QY 661 CAGCACACCGCAGAGACGACGAGCAGCAAGAGAGAAACGAAGGTGGAAACATCTTC 720
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlySerMetLeu 220
QY 721 AGCGCTTCACCGCGAGTTCCTCGAAACAGCTTCACAGTTCACGACACAGACAGATAGT 780
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValValAspArgGlnIleValVal 240
QY 781 CAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACGTGAGGGA 840
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 ArgLysLeuGlnGlyGluAsnGluGluGluLysGlyAlaIleValThrValArgGly 260
QY 841 GGCCTCAGAACTTTCAGGCCA-----GATAGAAGAGACGTCGCCACCAAGAGAG 891
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 GlyLeuSerValIleSerProThrGluGluArgGlnArgProGluGluGlu 280
QY 892 GAATACGATGAAGATCAATATCAATACATGAAGAGAGTAGAAGCGTGGCAGGGAAGC 951
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 295
QY 952 AGAGCAGCGGGGAATGGTATTGAAGAGACGATCTCCACCGCAAGTGCTAAAGAACAT 1011
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 Arg-----AsnGlyIleAspGluThrIleCysThrMetGlyLeuArgHisAsnIle 312
QY 1012 GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAACTGCCAAC 1071
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 332
QY 1072 GATCTCAACTCTTAATACCTTAGTGGTTCGACCTAGTGTGCTGAATATGGAATCTCTAC 1131
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
333 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 352
QY 1132 AGGAATGCATTTGTTGCTCACTACAACCAACGACACACAGCATCATATATCGATTG 1191

```

```

Db 353 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 372
Qy 1192 AGGGACGGGCTCAGTGAAGTCTGTCAGACAGCAACGGCAACAGAGTGTACGACGAGGAG 1251
Db 373 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 392
Qy 1252 CTTCAAGAGGGTCACGTGCTGTGGTCCACAGACTTCCGCGTCCGTGGAAAGTCCAG 1311
Db 393 LeuGlnGluGlnValLeuThrValProGlnAsnPheAlaValAlaAlaArgSerGln 412
Qy 1312 AGCGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGCGCCAGCATAGCCAACTC 1371
Db 413 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 432
Qy 1372 GCCGGTGAAGAACTCCGTATAGATAACCTGCCGAGGAGGTGGTGCATAATTCATATGGC 1431
Db 433 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGlnGluValIleGlnGlnThrPheAsn 452
Qy 1432 CTCCAAAGGGAGCAGCGACGCTTAAGACACCAACCCCTTCAAGTCTCTCGTTCCA 1491
Db 453 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 472
Qy 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 473 ProLysGluSerGlnArgArgValValAla 482

```

RESULT 15

```

Q41702
ID Q41702 PRELIMINARY; PRT; 498 AA.
AC Q41702;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Legumin A precursor.
OS Vicia sativa (Spring vetch) (Tare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
OX NCBI_TaxID=3908;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RA Nong V., Becker C., Muentz K.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -1- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC EMBL; Z32835; CAA83677.1; -.
CC PIR; S44294; S44294.
CC HSP; P04776; IPIX.
CC GO; GO:0045735; F:nutrient reservoir activity; IEA.
CC InterPro; IPR006045; Cupin.
CC InterPro; IPR007113; Cupin.
CC InterPro; IPR011051; RmlC_Like_cupin.
CC InterPro; IPR006044; Seedstore_11s.
CC Pfam; PF00190; Cupin; 2.
CC PRINTS; PR00439; 11SGLOBULIN.
CC PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Seed storage protein; Signal; Storage protein.
FT SIGNAL
FT CHAIN 1 21
FT CHAIN 22 498
SQ SEQUENCE 498 AA; 56564 MW; 3E75D87DB9A99699 CRC64;

```

Alignment Scores:

```

Pred. No.: 9.77e-94 Length: 498
Score: 1457.50 Matches: 289
Percent Similarity: 68.88% Conservative: 74
Best Local Similarity: 54.84% Mismatches: 93
Query Match: 54.16% Indels: 71
DB: 2 Gaps: 8

```

```

US-10-728-323-3 (1-1524) x Q41702 (1-498)
Qy 1 CGGCAGCAACCGGAGGAGAACGGTGCAGTTCAGCGCCTCAATGCGCAGAGACCTTGAC 60
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgIleAsnAlaLeuGluProAsp 42
Qy 61 AATCGCATTCGAATCAGAGGGCGGTTCATTGAGACTTTGGAACCCCAACAACAGGAGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnArgGlnPhe 62
Qy 121 GAATGCGCGGCTCGCCCTCTCTCGTTAGTCTCCGCGCAGCAACGCCCTTCGTAGGCCT 180
Db 63 ArgCysAlaArgValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro 82
Qy 181 TTCTACTCCAATCTCCCCAGGAGATCTTTCATCAGCAAGAAAGGGGATATCTTTGGGTG 240
Db 83 TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet 102
Qy 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCTCACACAAAGTCTCGATCTCAG 300
Db 103 ValPheProGlyCysProGluThrHisGluGluPro-----GlnGln 116
Qy 301 TCCCAAGACCAACCAAGACGCTCTCCAGGAGAGAGACCAAGCCCAACAGCAACGAGTAGT 360
Db 117 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspSer 128
Qy 361 CACCAGAGGTGCACCGTTCGATGAGGTGATCTCATTTGCAGTTCCTCCACGGTGTGCT 420
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleAla 148
Qy 421 TTCTGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspThrGly 168
Qy 481 AACAAACGACCAACGAGTTCATGATTCCTCCAGGAGATTCATTTGGTGGGAACACGGAG 540
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
Qy 541 CAAGAGTTCCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db 189 GlnGluPheLeuArgTyrGlnHisGln----- 197
Qy 601 AGCCCATACAGCCCGCAAAAGTACGCTAGACAAGAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 197 ----- 197
Qy 661 CAGCACAGCGCAGAGAACGAGCAGCAGCAAGAGAGAAAGAAACGAGTGGAAACATCTTC 720
Db 198 -----GlnGlyLysGlnGlnAspAsnAspGlyAsnAsnIlePhe 212
Qy 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTCCACGACACAGATAGTG 780
Db 213 SerGlyPheLysArgAspPheLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 231
Qy 781 CAAACCTTAAGCGCAGACCGGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGGA 840
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGlyGlyAlaIleValLysValLysGly 251
Qy 841 GGCCTCAGAATCTTGAGCCACAGATAAAG-----AGACGT 876
Db 252 GlyLeuSerIleIleAlaProGluArgGlnAlaArgHisGluArgGlySerArgGln 271
Qy 877 GCGCAGCAAGAAGAGGAATACGATGAA-----GAT 906
Db 272 GluGluAspGluAspGluLysGluGluArgGlnProSerHisHisLysSerArgArgAsp 291
Qy 907 GAATATGAATACGATGAAGAGGATAGA-----AGCGGTGGCAGGGAAGCAGAGGAGG 960
Db 292 GluAspGluAspAspLysGluLysArgHisSerGlnLysGlyGlnSerArgArgGlnGly 311
Qy 961 GGGATGGTATTCAAGAGACGATCTCCAGCGCAAGTCTCTAAAAGAACATTGGTAGAAC 1020
Db ----- 1020

```


Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	811	30.1	492	4	US-09-462-720-2	Sequence 2, Appli
2	293	10.9	141	4	US-09-645-593-12	Sequence 12, Appl
3	227.5	8.5	85	4	US-09-645-593-10	Sequence 10, Appl
4	215	8.0	165	4	US-09-645-593-11	Sequence 11, Appl
5	184.5	6.9	489	4	US-09-424-283-3	Sequence 3, Appli
6	173.5	6.4	96	4	US-09-645-593-9	Sequence 9, Appli
7	172.5	6.4	626	4	US-09-106-872A-4	Sequence 4, Appli
8	167	6.2	524	4	US-09-424-283-1	Sequence 1, Appli
9	160	5.9	448	4	US-09-323-195A-18	Sequence 18, Appl
10	158	5.9	523	4	US-09-323-195A-17	Sequence 17, Appl
11	157.5	5.9	335	4	US-09-106-872A-17	Sequence 17, Appl
12	156.5	5.8	587	1	US-07-955-905A-23	Sequence 23, Appl

QY	52	AGACCTGACAATCGCATTTGAATCAGAGCGCGGTACATTGAGACTTGGAAACCCCAACAC	111
DB	45	GluproSerPheArgPheProSerGluAlaGlyLeuThrGluPheThrAspSerAsn	64
QY	112	CAGAGATTGCAATCGCGCGCGTGCCTCTCTCGTTAGTCTCCGCCGCAACGCCCTT	171
DB	65	ProGluPheGlyCysAlaGlyValGluPheGluArgenThrValGlnProIysGlyLeu	84
QY	172	CGTAGCGCTTTCTACTCCAATGCCTCCACGAGATCTTCATCCAGCAAGGAGGGATAC	231
DB	85	ArgLeuProHisIstYrSerAsnValProIysPheValTyrValValGluGlyThrGlyVal	104
QY	232	TTTGGTTGATATTCCTCGTTTCTTAGACACTATGAAGAGCTCACACAAGAGTCGT	291
DB	105	GlnGlyThrValIleProGlyCysAlaGluThrPheGlu-----SerGlnGlyGlu	121
QY	292	CGA-----TCTCAGTCCCAAGACCACCAAGAGCTCTCCAAGGAGAGACCA-----	339
DB	122	SerPheTrpGlyGlnGlnGlnProGlyIysGlyGlnGluGlyGlnGlnGlySer	141
QY	340	-----AGCCAAACAGCAACGA-----GATAGTCCACAGAAGGTCCACCGT	378
DB	142	LysGlyGlyGlnGluGlyArgArgGlnArgPheProAspArgHisGlnIysLeuArgArg	161
QY	379	TTCCATGAGGTTGATCTCATTTGCAGTTCCTCCACCGGTGTGCTTCTTGCTCTACAACGAC	438
DB	162	PheGlnIysGlyAspValLeuLeuLeuProGlyPheThrGlnTrpThrTyrAsnAsp	181
QY	439	CACGACACTGATGTTGTGCTGTTCTTCTTACTCACACCAACAACGACCAACAGCTT	498
DB	182	GlyAspValProLeuValThrValAlaLeuLeuAspValAlaAsnGluAlaAsnGlnLeu	201
QY	499	GATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACACGGAGCAAGAGTTCTTTAAGGTAC	558
DB	202	AspLeuGlnSerLysPhePheLeuAlaGlyAsnProGlnGlnGlyGlyGlyGlyGlu	221
QY	559	CAGCAACAACAGCAGCAAGCAGCAGCAAGAGAGCTTTACCATATAGCCCATACAGCCGCA	618
DB	222	GlyHisGlnGlyGlnGlnGlnHisArg-----	231
QY	619	AGTCAGCCTACACAAGAGCGGTGAATTTAGCCTCGAGACAGCAGCAGCCGCAAGAA	678
DB	231	-----	231
QY	679	CGAGCAGGACAAGAAGAAAAAGAGGTGGAACATCTTCAGCGCTTCACGCCGAG	738
DB	232	-----AsnIlePheSerGlyPheAspGln	240
QY	739	TTCTTGGAACAGCCTTCCAGTTGCAGACACAGATAGTGTGCAAAACCTTAAGAGCGCAG	798
DB	241	LeuLeuAlaAspAlaPheAsnVal--AspLeuIysIleIleGlnIysLeuLysGly---	258
QY	799	ACCGAGCTGAAGAAGAGCGGAGCCATTGTGTGACGTGAGGGA-----GGCTCAGAACTTGT	855
DB	259	---ProIysAspGlnArgIysSer-----ThrValArgAlaGluLysLeuGlnLeuPhe	275
QY	856	AGCCAGATAGAAAGAGACGTCGCCGCAAGAGGAATACGATGAAGATGATATGAA	915
DB	276	LeuProGluTrpSerGluGlnValGlnGlnProGlnGln-----	288
QY	916	TACGATGAAGAGGATAGAAGCGTGGCAGGGGAACGAGCAGCGGGGAATGGTATTGAA	975
DB	289	---GlnGlnGlnGlnGlnHisGlyValGlyArgGlyTrpArgSerAsnGlyLeuGlu	307
QY	976	GAGACGATCTGCACGCGAAGTGCTAAAGAACATTGGTAGAAACAGATCCCTTCAGCATC	1035
DB	308	GluThrLeuCysThrValIysLeuSerGluAsnIleGlyLeuProGlnGluAlaAspVal	327
QY	1036	TACAACCTCAAGCTGTTCACTCAAACCTGCCAACGATCTCAACCTTCTTAATACTTAGG	1095
DB	328	PheAsnProArgAlaGlyArgIleThrValAsnSerGlnIysIleProIleLeuSer	347
QY	1096	TGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAATGATGTTTGTTCGCTCAC	1155

Db	348	SerLeuGlnLeuSerAlaGluArgGlyPheLeuTySerAsnAlaIlePheAlaProHis	367
Qy	1156	TACAACACCAACGACACACAGCATATATCGATTGAGGGGACGGGTCCACGTGCAAGTC	1215
Db	368	TrpAsnIleAsnAlaHisAsnAlaLeuTyValIleArgGlyAsnAlaArgIleGlnVal	387
Qy	1216	GTGCACAGCAACGGCACACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTGTGTGTG	1275
Db	388	ValAspHisIysGlyAsnLysValPheAspGluValLysGlnGlyGlnLeuIleIle	407
Qy	1276	GTGCCACAGAACTTCGCGTCGCTGGGAAAGTCCACAGAGCAGAACTTCGAATACGTGGCA	1335
Db	408	ValProGlnTyrPheAlaValIleLysLysAlaGlyAsnGlnGlyPheGluTyrValAla	427
Qy	1336	TTCAAGACACACTCAAGGCCGACATAGCCAACTCGCCGGTGAACACTCCGTCATAGAT	1395
Db	428	PheLysThrAsnAspAsnAlaMetIleAsnProLeuValGlyArgLeuSerAlaPheArg	447
Qy	1396	AACTCTCGGAGGAGGTGGTTGCAAAATTCATATATGGCTTCCAAAGGGAGCAGGCAAGCAG	1455
Db	448	AlaIleProGluGluValLeuArgSerSerPheGlnIleSerSerGluGluAlaGluGlu	467
Qy	1456	CTTAAGAACCAACACCCCTTCAAGTCTTCGTTCACCGTCTCAGCAGTCT	1506
Db	468	LeuLysTyrGlyArgGlnGluArgLeuLeuLeu-----SerGluGlnSer	482
RESULT 2			
US-09-645-593-12			
; Sequence 12, Application US/09645593			
; Patent No. 6777591			
; GENERAL INFORMATION:			
; APPLICANT: Chaudhary, Sarita			
; APPLICANT: van Rooijen, Gijs			
; APPLICANT: Moloney, Maurice			
; APPLICANT: Singh, Surinder			
; TITLE OF INVENTION: Flax Seed Specific Promoters			
; FILE REFERENCE: 9369-151			
; CURRENT APPLICATION NUMBER: US/09/645,593			
; CURRENT FILING DATE: 2000-08-25			
; PRIOR APPLICATION NUMBER: US 60/151,044			
; PRIOR FILING DATE: 1999-08-27			
; PRIOR APPLICATION NUMBER: US 60/161,722			
; PRIOR FILING DATE: 1999-10-27			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 12			
; LENGTH: 141			
; TYPE: PRT			
; ORGANISM: Linum usitatissimum			
US-09-645-593-12			
Alignment Scores:			
Pred. No.:	1,21e-19	Length:	141
Score:	293.00	Matches:	55
Percent Similarity:	72.73%	Conservative:	25
Best Local Similarity:	50.00%	Mismatches:	30
Query Match:	10.89%	Indels:	0
DB:	4	Gaps:	0
US-10-728-323-3 (1-1524) x US-09-645-593-12 (1-141)			
Qy	1138	GCATTGTTTTCGCTCATTACACACCAACGACACAGCATCATATATCGATTGAGGGA	1197
Db	2	AlaIleArgLeuProHisTrpAsnIleAsnAlaHisSerIleValIyrAlaIleArgGly	21
Qy	1198	CGGGCTCACGTGCAAGTCGTGGGACCAACGGCAACAGAGTGTACGACGAGGAGCTTCAA	1257
Db	22	GlnAlaArgValGlnIleValAsnGluGluGlyAsnSerValPheAspGlyValLeuGln	41
Qy	1258	GAGGGTCACTGTTGTGTGGTCCACAACTTCGCGCTCGCTCGAAGTCCCGAGAGCGAG	1317
Db	42	GluGlyGlnValThrValProGlnAsnPheAlaValValLysArgSerGlnSerGlu	61

```
QY 1318 AACTCGAATACGTGGCACTCAAGACAGACTCAAGGCCAGCATAGCAACCTCGCGGT 1377
Db 62 ArgPheGluTrpValAlaPheLysThrAsnAspAsnAlaMetValAsnSerLeuAlaGly 81
QY 1378 GAAACCTCCGTCATAGATAACCTCGCGAGGAGGTGTGCAAAATTCATATGGCTCCAA 1437
Db 82 ArgThrSerAlaValArgAlaIleProAlaAspValLeuAlaAsnAlaTrpArgValSer 101
QY 1438 AGGAGCAGGCAAGCGCAGCTTAAGAACAA 1467
Db 102 ProGluGluAlaArgArgValLysPheAsn 111

RESULT 3
US-09-645-593-10
; Sequence 10, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijs
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (78)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (82)
; OTHER INFORMATION: n is any amino acid
US-09-645-593-10

Alignment Scores:
Pred. No.: 1,9e-13 Length: 85
Score: 227.50 Matches: 46
Percent Similarity: 61.22% Conservative: 14
Best Local Similarity: 46.94% Mismatches: 25
Query Match: 8.45% Indels: 13
DB: 4 Gaps: 2

US-10-728-323-3 (1-1524) x US-09-645-593-10 (1-85)
QY 220 GGAAGGGGATACCTTGGGTGATATTCCTGTGCTAGACACTATGAAGACCTCAC 279
Db 1 GlyArgGlyValThrGlyIleMetPhePro**CysProGluThrPheGluGluSerGln 20
QY 280 ACAACAGGTGCTGCATCTCACTCCAAAGACCACCAAGACGTCTCCAAAGGAGAACCAA 339
Db 21 GlnGlnGlyGlnGlnGlyGln-----GlnGlySerSerGln 32
QY 340 AGCCAAACAGCAACGAGATAGTACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCAT 399
Db 33 -----AspGlnHisGlnLysIleArgArgPheArgGluGlyAspValIle 47
```

```
QY 400 GCAGTTCACCGCTGCTTCTGGCTCTACAAACGACACCACTGATGTGTGCT 459
Db 48 AlaValProAlaGlyValAlaHisTrpSerTrpAsnAspGlyAsnGluProValMetAla 67
QY 460 GTTCTCTTACTCACACCAACAAACAGCAACAGCTTGCATCAGTTCCCCAGG 513
Db 68 IleValValHisAspThrSerSerHisLeuAsnGlnLeuAspAsnProArg 85

RESULT 4
US-09-645-593-11
; Sequence 11, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijs
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Linum usitatissimum
US-09-645-593-11

Alignment Scores:
Pred. No.: 4,29e-12 Length: 165
Score: 215.00 Matches: 62
Percent Similarity: 43.06% Conservative: 28
Best Local Similarity: 29.67% Mismatches: 69
Query Match: 7.99% Indels: 50
DB: 4 Gaps: 6

US-10-728-323-3 (1-1524) x US-09-645-593-11 (1-165)
QY 517 TTCATTTGCTGGGAACACGGAGCAAGAGTCTTAAAGTACCAGCAACAAACGACAA 576
Db 2 PheTyrLeuAlaGlyAsnProArgAspGluPheGluGlnSerGlnGlnGlyArgLeu 21
QY 577 AGCAGACGAGAAGCTTACCATATAGCCCAAGCCGCAAGTCCAGCTAGACAAGAA 636
Db 22 SerArgGlyGluSer----- 26
QY 637 GAGCGTGAATTTAGCCCTCGAGGACACAGCCGCGAGAGACGAGCAGGACAAAGAA 696
Db 27 -----GluGlyGlyArgGlyArgGluProLeuGlnProAlaThr 40
QY 697 GAAAACGAGGTGGAAACATCTTCAGCGCTTCACCCGAGTCTCTGGAACAGCTTC 756
Db 41 ThrSerSer-----CysGlyIleAspSerLysLeuIleAlaGluAlaPhe 55
QY 757 CAGGTTGACGACACAGATAGTGCAAAACCTTAAGAGCGGACCCAGAGTGAAGAAGAG 816
Db 56 AsnValAspGlu---AsnValAlaArgArgLeu-----GlnSerGluAsnAspAsnArg 72
QY 817 GGAGCCATTGTGACAGTGAAGGGGAGCCCTCAGAATCTTGAGCCCAAGTGAAGACAGCT 876
Db 73 GlyGlnIleValArgValGluGlyGluLeuAspIleValArgProThrSerIleGln 92
QY 877 GCCGCAAGAGAAGAGATACCATGATGAATATATGATATGATGATGATGATGATGATG 936
Db 93 GluGluSerGlnGluGln----- 98
QY 937 CGTGGCAGGGGACGAGCAGCGAGCGGG-----AATGGTATTGAAGACAGATC 984
```

Db 99 -----GlyGlyArgGlyGlyArgTyrTyrSerAsnGlyValGluGluThrPhe 115
Qy 985 TGCACCCCAAGTGTCTAAAAAGAACATTGGTGTAGAACAGATCCCTCGACATCTTACAAACCT 1044
Db 116 CysSerMetArgLeuIleGluAsnIleGlyAspProSerArgAlaAspIlePheThrPro 135
Qy 1045 CAAGCTGGTCTCACTCAAACTGCCAACGATCTCAACCTTCTAATCTTAGTGGCTTGGGA 1104
Db 136 GluAlaGlyArgValArgSerLeuAsnSerHisAsnLeuProValLeuGluThrPheGln 155
Qy 1105 CCTAGTCTGAATATGGAATCTCTAC 1131
Db 156 LeuSerAlaGluArgGlyValLeuTyr 164
RESULT 5
US-09-424-283-3
; Sequence 3, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-3
Alignment Scores:
Pred. No.: 6,51e-09 Length: 489
Score: 184.50 Matches: 99
Percent Similarity: 34.19% Conservative: 87
Best Local Similarity: 18.20% Mismatches: 187
Query Match: 6.86% Indels: 171
DB: 4 Gaps: 17
US-10-728-323-3 (1-1524) x US-09-424-283-3 (1-489)
Qy 1 CGGCAGCAACGGGAGGAGACGGTCCAGTTCACGGCGCTCAATGGCGAGACCT--- 57
Db 80 ArgGluLysGluGluGluHisGlnGluGlnHisGluGluGluGluAspGluAsnProTyr 99
Qy 58 -----GACAATCGCATTGAATCAGAGCGCGGTACATTGAG 93
Db 100 ValPheGluGluAspLysAspPheSerThrArgValGluThrGluGlySerIleArg 119
Qy 94 ACTTGGAAACCCCAACACAGGAGTTCGAA---TGGCGCGCTTCGCCCTCTCTCGCTTA 150
Db 120 ValLeuLysLysPheThrGluLysSerLysLeuGlnGlyIleGluAsnPheArgLeu 139
Qy 151 GTCTC-----CGCCCAACGGCTTCGTAGGCTTCTACTCCAATGCTCCCCAG 201
Db 140 AlaIleLeuGluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluVal 159
Qy 202 GAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTGGATATTCCTGCTGTTGCTCTAGA 261
Db 160 ValLeuPheAsnIleLysGlyArgAlaValLeuGlyLeuVal----- 173
Qy 262 CACTATGAAGAGCTCACACAAAGGTGTCGATCTCAGTCCCAACCAACCAACGACGT 321
Db 173 ----- 173
Qy 322 CTCCAGGAGAGACCAACCAACAGACGAGATAGTCACCAAGGTGACCGCTTC 381
Db 174 -----ArgGluSerGluThrGluLysIleThrLeu 183

Qy 382 GATGAGGTGATCTCATTCAGTTCACACCGGTGTTGCTTTCTGCTCTACACGACCAC 441
Db 184 GluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArgAsp 203
Qy 442 GACACTGATGTTGCTGCTCTTCTTACTGTACCAACCAACACGACGACGACGCTTGAT 501
Db 204 GluAsnGluLysLeuLeuAlaMetLeu----- 213
Qy 502 CAGTTCCTCCAGGAGATTCAATTGGCTGGGAACCGGACGAGAGTTCCTTAAGGTACCAG 561
Db 213 ----- 213
Qy 562 CAACAAAGCAGACAAAGCAGACAGAGTTCACATATAGCCCATATACAGCCGCAAAAGT 621
Db 214 -----HisIleProValSer 218
Qy 622 CAGCCTAGACAAGAGAGCGTGAAATTAGCCTCGAGGACAGACAGCGCGCAGAGAACA 681
Db 219 ThrProGlyLysPheGluGluPheGlyProGlyGlyArg----- 232
Qy 682 GCAGGACAAAGAAAGAAACCAAGGTGGAACATCTTCAGCGGCTTCACGCGGAGTTC 741
Db 233 -----AspProGluSerValLeuSerAlaPheSerTrpAsnVal 245
Qy 742 CTGGAACAAAGCCTTCAGGTTGACGACACAGATAGTCAAAACCTTAAGAGCGCGAGACC 801
Db 246 LeuGlnAlaAlaLeuGlnThrPro-----LysGlyLysLeu 257
Qy 802 GAG-----AGTGAAGAAGAGGAGCGCATTTGTACAGTG---AGGGAGGCGCTC 846
Db 258 GluArgLeuPheAsnGlnGlnAsnGluGlySerIlePheLysIleSerArgGluArgVal 277
Qy 847 AGAATCTTCAGCCCGACATAGAAAGAGACGTGCCGACGAAAGAGAGAAATACGATGAAGAT 906
Db 278 ArgAlaLeuAlaProThrLysLys-----SerSer 287
Qy 907 GAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAAGCAGAGCGAGGGGAAT 966
Db 288 TrpTrpProPheGlyGlySerLys----- 296
Qy 967 GGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACATTGGTGTAGAACAGATCC 1026
Db 297 -----AlaGlnPheAsnIlePheSerLysArg--- 305
Qy 1027 CCTGACATCTACAAACCTCAAGCTGTTTCACTCAAAACCTGCCAACGATCTCAACCTCTA 1086
Db 306 ProThrPheSerAsnGlyTyrGlyArgLeuThrGluValGlyProAspAspGluLysSer 325
Qy 1087 ATACTTAGTGTGCTTGGACCTAGTGTGAATATGAAATCTCTACAGGAATGCATTGTTT 1146
Db 326 TrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSerMetSer 345
Qy 1147 GTCGCTCACTACACACCAACCGCACACGATCATATATCGATTGAGGGGACGGGCTCAC 1206
Db 346 ThrIleHisTyrAsnSerHisAlaThrIleAlaLeuValMetAspGlyArgGlyHis 365
Qy 1207 GTGCAAGTC-----GTGGACACGACGCGCAACAGAGGTGTACACGAG----- 1248
Db 366 LeuGlnIleSerCysProHisMetSerSerArgSerAspSerLysHisAspLysSerSer 385
Qy 1249 -----GAGCTTCAAGAGGGTACGCTGCTTGTGTGGTCCACAG 1284
Db 386 ProSerTyrHisArgIleSerAlaAspLeuLysProGlyMetValPheValProPro 405
Qy 1285 AACTTCGCC---GTGCTGGAAGTCCCGACGCGAGAACTTCGAATACGTGCATTCAG 1341
Db 406 GlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuIleCysPheGlu 425
Qy 1342 ACAGACTCAAGCCCGCCAGCATAGCC---AACCTCGCGGTGAAACTCCGCTCAT--- 1392
Db 426 ValAsnValArgAspAsnLysLysPheThrPheAlaGlyLysAspAsnIleValSerSer 445

US-10-728-323-3 (1-1524) x US-09-323-195A-18 (1-448)	
QY 4 CAGCAACCGGAGGAGACGGTGCAGTTCAGCGCCTCAATCGCAGAGACCTGACAAT 63	
Db 46 GluGluArgGluGluAsnProTyrValPhe-----HisSerAspSerPheArgThr 62	
QY 64 CGCATTTGAATCAGAGGCGGTTCATTAGACTTGGAAACCCCAACACAGGAGTTTCGAA 123	
Db 63 ArgAlaSerSerGluAlaGlyGluIleArgAlaLeu---ProAsnPheGlyGluValSer 81	
QY 124 -----TGCCTGGGTGCGCCTCTCTCGTTAGTC-----CTCCGCCCAAGGCC 168	
Db 82 GluLeuLeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThr 101	
QY 169 CTTTCGTAGGCTTTCTACTCCAATGCTCCCGAGGATCTTCATCCAGCAAGCAAGCGGA 228	
Db 102 ValMetLeuProHisTyrIleAspAlaThrTrpIleLeuTyrValThrArgGlyArgGly 121	
QY 229 TACTTTGGGTGTATATTCCTGTTGCTGTAGACACTATGAAGAGCTCACACACAAGGT 288	
Db 122 TyrIleAlaTyrVal-----126	
QY 289 CGTCGATCTCAGTCCCAAGACCAACAGACGTCTCCAAGGAGAGACCAAGCCAACAG 348	
Db 127 -----HisGlnAsnGluLeuValLysArgLysLeu-----136	
QY 349 CAACGAGATAGTACCAGAGGTGCACCGTTTCGATCAGGCTCATCTCATTGCAGTCCC 408	
Db 137 -----GluGlyAspValPheGlyValPro 145	
QY 409 ACCGGTGTCTTTGGCTCTACAACGACACGACTGATGTTGTGTGTTCTCTTT 468	
Db 146 SerGlyHisThrPheTyrLeuValAsnAspHisAsnThrLeuArgIleAla---164	
QY 469 ACTGACACCAACAAACAGCAACACGACTTGATCAGTTCCCGAGGAGATTCAAATTTGGCT 528	
Db 164 -----164	
QY 529 GGGAAACCGGAGCAGAGTTCTTAAGTACCAGCAACAAAGCAGACAAAGCAGCAGAA 588	
Db 164 -----164	
QY 589 AGCTTACCATATAGCCCATACAGCCGCAAGTCAGCTAGACAAGAAGCGTGAATTT 648	
Db 165 -----SerLeuValArgProValSerThrValArgGlyGluTyrGlnProPhe 180	
QY 649 AGCCCTCGAGGACGACCCGCGAGAGCAGGAGCAAGAAAGAAAGCAAGGT 708	
Db 181 TyrValAlaGly-----GlyArgAsnProGln-----189	
QY 709 GGAACATCTTCAGCGCTTCACGCGGAGTTCTCTGCAACAGCCTTCAGGTTGACGAC 768	
Db 190 -----ThrValTyrSerAlaPheSerAspValLeuGluAlaAlaPheAsnThrAsnVal 208	
QY 769 AGACAGATAGTCAAAACCTTAAGAGGCGAGACCGCAGAGTGAAGAAGAGGAGCCATTGTG 828	
Db 209 GlnGlnLeuGluArgIlePheGlyGly-----HisLysSerGlyValIleIle 224	
QY 829 ACAGTGAGGGAGCGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGCAAGAA 888	
Db 225 -----HisAlaAsnGluGlu 229	
QY 889 GAGGAATACGATGAGTGAATGATATGATACGATGAGAGGATAGAGGCTGCAGGGGA 948	
Db 230 -----GlnIleArgGluMetMetArgLysArgGly-----Phe 240	
QY 949 AGCAGAGCGCAGGGGAATGTTTGAAGAGACGATCTGCACCCGCAAGTGTCTAAAGAAAC 1008	
Db 241 SerAlaGlySerMetSerAlaProGluHis-----ProLysProPhe 254	
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACAACCCCTCAGCTGGTTCATCTCAAACTGCC 1068	

255	AsnLeuArgAsnGlnLysProAspPheGluAsn---GluAsnGlyArgPheThrIleAla	273
1069	AACGATCTCAACTTCTTAATAGTGGCTTGGACCTTAGTGGCTGAATATGGAATCTC	1128
Db	274 GlyProLysAsnTyrProPheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeu	293
QY	1129 TACAGGAATGCATTTGTTGTCGCTCACTAACACCAACACACAGCATCATATATCGA	1188
Db	294 AsnProGlySerMetThrAlaProSerLeuAsnSerLysSerThrSerIleGlyIleVal	313
QY	1189 TTGAGGGGACGGCTCACGTGCAAGTCGTG-----1218	
Db	314 ThrAsnGlyGluGlyArgIleGluMetAlaCysProHisLeuGlyGlnHisGlyTrpSer	333
QY	1219 -----GACACCAACGGCAACAGAGTGTACACGAGGAGCTT	1254
Db	334 SerProArgGluArgGlyAspGlnAspIleThrTyrGlnArgValTrp---AlaLysLeu	352
QY	1255 CAAGAGGTCACGTGCTTGTGTGCCACAGAACTTGGCCGTCGCTGGA---AAGTCCCA	1311
Db	353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr	372
QY	1312 AGCGAGACTTCGATATACGTGCATTCAAGACAGACTCAAGGCCACGATAGCCAAC---	1368
Db	373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe	392
QY	1369 CTCGCGGTGAAACTCCGTCATAGTAACTCCGCGAGAGGTGGTTGCAAAATTCATAT	1428
Db	393 LeuAlaGlyLysAsnValLeuAsnThrLeuGluArgGluIle-----407	
QY	1429 GGCCTCCAAAGGAGCAGCGCAAGCGAGCTTAAGAACAACAACCC 1473	
Db	408 -----ArgGlnLeuSerPheAsnValPro 415	
RESULT 10		
US-09-323-195A-17		
; Sequence 17, Application US/09323195A		
; Patent No. 6462257		
; GENERAL INFORMATION:		
; APPLICANT: Pullman, Gerald		
; APPLICANT: Cairney, John		
; APPLICANT: Ferreira, Ranjan		
; TITLE OF INVENTION: VITILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND		
; TITLE OF INVENTION: METHODS OF USING THE SAME		
; FILE REFERENCE: IPST0009		
; CURRENT APPLICATION NUMBER: US/09/323,195A		
; CURRENT FILING DATE: 1999-06-01		
; NUMBER OF SEQ ID NOS: 19		
; SOFTWARE: Patent In Ver. 2.0		
; SEQ ID NO 17		
; LENGTH: 523		
; TYPE: PRT		
; ORGANISM: Pinus taeda		
US-09-323-195A-17		
Alignment Scores:		
Pred. No.:	2,41e-06	Length: 523
Score:	158.00	Matches: 107
Percent Similarity:	32.67%	Conservative: 74
Best Local Similarity:	19.31%	Mismatches: 190
Query Match:	5.87%	Indels: 183
DB:	4	Gaps: 20
US-10-728-323-3 (1-1524) x US-09-323-195A-17 (1-523)		
QY 10 CCGGAGGAGAACGGTGCAGTTCAGCGCCTCAATCGCAGAGACCTGACAAAT-----	63	
Db 37 ProGluAspHisGlyArgGlyHisGlnArgArgGluGluArgGluGluAsnProTyr	56	
QY 64 -----CGCATTTGAATCAGAGGCGGTTCATCATTTGAGACT	96	
Db 57 ValPheHisSerAspArgPheArgMetArgAlaSerSerAspAlaGlyGluIleArgAla	76	

```
QY 97 TGAACCCCAACACAGGAGTTCGAA-----TGC GCGCGGTGCGCTCTCTCGCTTA 150
Db 77 Leu---ProAsnGlyGluLeuSerGluLeuGluGlyLeuSerLysTyrArgVal 95
QY 151 GTC-----CTCCGCGCAACGCCCTTCGTAGGCTTCTACTCAATGCTCCCCAG 201
Db 96 ThrCysIleGluMetArgProAsnThrValMetLeuProHisTyrLeuAspAlaThrTrp 115
QY 202 GAGATCTTCATCCAGCAAGAGGAGTACTTTGGTGTATATTCCTCGTGTTCCTAGA 261
Db 116 IleLeuTyrValThrGlyGlyArgGlyTyrIleAlaTyrVal-----129
QY 262 CACTATGAAGAGCTCACACACAGGTGCTCGATCTCAGTCCCAAGACACCAAGACGT 321
Db 130 -----HisGlnAsnGluLeuValLysArgLys 138
QY 322 CTCAAGAGGAGAACCAAGCCAAAGCAGACAGATAGTCACCAAGAGGTGCACCGTTTC 381
Db 139 Leu-----139
QY 382 GATGAGGTGATCTCATTGTCAGTTCCACCGGTGTTGCTTTCTGGCTCTCAACGACAC 441
Db 140 GluGluGlyAspValPheGlyValProSerGlyHisThrPheTyrLeuValAsnAsp 159
QY 442 GACACTGATGTTGCTGTTCTCTTACTGACACCAACACACAGCAACACCGACTTGAT 501
Db 160 AspHisAsnSerLeuArgIleThr-----167
QY 502 CAGTTCCCGAGGAGATTCAATTTGGCTGGGAACACAGGCAAGAGTTCTTAAGTACCAG 561
Db 168 -----SerLeuLeuArgThrVal 173
QY 562 CAACAAGCAGACAAAGCAGACGAGAGAGTTACATATAGCCCATACAGCCCGCAAGT 621
Db 174 SerThrMetArgGlyGlu-----TyrGluProTyr-----183
QY 622 CAGCTAGACAGAGAGCGGTGAATTTAGCCCTCGAGGACAGACACCGCAGAGACGA 681
Db 184 -----TyrValAla 186
QY 682 GCAGGACAGAGAGAGAGAGAGGTGGAAACATCTTCAGCGGCTTCAGCGCGAGTTC 741
Db 187 GlyGlyArgAsnProGlu-----ThrValTyrSerAlaPheSerAspVal 202
QY 742 CTGGAACAGCCTTCCAGTTGACGACAGACAGATAGTCAAACTTAAAGGCGAGACC 801
Db 203 LeuGluAlaAlaPheAsnThrAsn-----210
QY 802 GAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGAGGCGCTCAGAACTTCAGCCCA 861
Db 211 -----ValIleGluAlaArgThrHisPheProVal-----220
QY 862 GATAGAAAGAGACGTGCCGAGAGAGAGAAATACGATGAAGATGAATATGAATACGAT 921
Db 221 -----HisIleGluArgGluSerTyrSerMetAla 230
QY 922 GAAGAGATAGAGGCGTGGCAGGGAAGCAGAGCGAGGGGAATGTGTATTGAAGAGCG 981
Db 231 AsnGluGluGlnIleArg---GluMetLeuArgLysArgGlyPheSerAlaGluSerMet 249
QY 982 ATCTGCACCGCAAGTGTCTAAAGAACATTTGGTAGAACAGATCCCTCGCATCTACAAC 1041
Db 250 SerAlaSerGluHisProLysProPheAsnLeuArgAsnGlnLysProAspPheGluAsn 269
QY 1042 CCTCAAGTGGTTCACCTCAAACTGCCAACGATCTCAACCTCTTAATCTAGTGGCTT 1101
Db 270 AspAsnGlyArgPheThrArgAlaGlyProAsnGluAsnProLeu---LeuAspAlaVal 288
QY 1102 GGACCTAGTCTGAATATGAAATCTCTACAGGAATGCATTTGTTGCGTCTCACTACAAC 1161
Db 289 AspValThrAlaGlyPheGlyValLeuAsnProGlyThrMetThrAlaProSerHisAsn 308
QY 1162 ACCAACGCACACATCATATATCGATTGAGGGGCGGCTCACGTCGCAAGTCGTG---1218
```

```
Db 309 ThrLysAlaThrSerIleAlaIleValThrGlnGlyGluGlyArgIleGluMeAlaCys 328
QY 1219 -----GACAGCAAC 1227
Db 329 ProHisLeuGlyGlnHisGlyTyrSerSerArgArgGluLysGlyAspGlnGluIleAsn 348
QY 1228 GGCACACAGAGTGTACGACGAGGAGCTTCAAGAGGTCACGCTGCTTGTGGTCCACAGAAC 1287
Db 349 TyrGlnArgVal---ArgAlaArgLeuArgThrGlyThrValTyrValValProAlaGly 367
QY 1288 TTCGCGCTGCTCGAAAGTCCACAGAGCGAG---AACTTCGAATACGTGGCATTCAAGACA 1344
Db 368 HisProIleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTrpPheAspIle 387
QY 1345 GACTCAAGGCCCATAGTACCAAC---CTCCGCGGTGAAACCTCCGTCATAGATAACCTG 1401
Db 388 AsnThrSerGlyAsnGluArgGlnPheLeuAlaGlyLysTyrAsnValLeuGlnThrLeu 407
QY 1402 CCGAGAGAGTGTGTCAAATTCATATGGCTCTCAAAAG---GACCAG 1446
Db 408 GluLysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluVal 427
QY 1447 GCAAGCGAGCTTAAGAAC---GAGGAGAGTGTGTCAAATTCATATGGCTCTCAAAAG---AACAAC 1470
Db 428 LeuArgArgGlnLysAspGlnValIleLeuArgGlyProGlnMetGlnArgArgAspGlu 447
QY 1471 CCCTCAAGTTCCTTCGTTCCACCGTCTCAGCAGTCTCTCCGA 1510
Db 448 -ProArgSerSerSerIleHisAlaIleIleAlaAlaArg 460

RESULT 11
US-09-106-872A-17
; Sequence 17, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina B.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-17

Alignment Scores:
Pred. No.: 2,14e-06 Length: 335
Score: 157.50 Matches: 55
Percent Similarity: 42.04% Conservative: 48
Best Local Similarity: 22.45% Mismatches: 81
Query Match: 5.85% Indels: 61
DB: 4 Gaps: 9

US-10-728-323-3 (1-1524) x US-09-106-872A-17 (1-335)
QY 31 TTCACGCGCTCAATCGCAGAGACCTGCAATCGCATTCGATTCAGAGCGGTACATT 90
Db 122 PheGlyLysLeuPheGluValLysProAspLysLys-----133
QY 91 GAGACTTGGNACCCCAACCAACAGGAGTTCGAATGCGCGCGCTCGCTCTCTCGCTTA 150
Db 91 GAGACTTGGNACCCCAACCAACAGGAGTTCGAATGCGCGCGCTCGCTCTCTCGCTTA 150
```

```
Db 134 -----AanProGlnLeuGlnAspLeuAsp-----MetMetLeuThrCysVal 147
Qy 151 GTCTCTCGCGCAACGCCCTTCGTAGGCTT---TTCTACTCCAATGCTCTCCAGGAGATC 207
Db 148 GluIleLysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleVal 167
Qy 208 TTCAATCAGCAAGAGGGGATACTTTGGGTGATATTCCTCTGTTGCTCTAGACACTAT 267
Db 168 ValValAsnLysGlyThrGlyAsnLeuVal----- 179
Qy 268 GAAGAGCTCACACACAGGTGCTGATCTCAGTCCCAAGACACCAAGACGCTCTCCAA 327
Db 180 -----AlaValArgLysGluGlnGlnArgGlyArgArgGluGlu 193
Qy 328 GGAGAGACCAAGCCCAACAGCAGATAGTACACAGAGGTGCACCGTTTC----- 381
Db 194 GluGluAspGluGluGluGluGluGlySerAsnArgGluValArgArgThrAla 213
Qy 382 -----GATGAGGTGATCTATTGTCAGTTCCACCGGTGTTCTTCTGCTCTACAAC 435
Db 214 ArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIleAsnAla 233
Qy 436 GACCACACACTGATGTTGTTGTTCTTCTTACTGACACCAACCAAC----- 486
Db 234 SerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePhe 253
Qy 487 -----GACAACAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
Db 254 LeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAspLeuAla 273
Qy 517 TTCAATTGCTGGACACCGGAGCAGAGCTTTTAAGTTACCAAGCAACCAAGCAGACAA 576
Db 274 PheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSerHisPhe 293
Qy 577 ACAGACAGCAAGACTTACCATATAGCCATAGCCGCAAGTACGCTAGACAGAA 636
Db 294 ValSerAlaGlnSerGlnSerGlnSerProSerProSerProGluLysGlu----- 309
Qy 637 GAGCGTGAATTAGCCCTCGAGGACAGCACAGCGCGCAGAGACGAGGAGCAAGAGAA 696
Db 310 -----SerPro-----GluLysGluAspGlnGluGlu 318
Qy 697 GAAACGAGAGTGA 711
Db 319 GluAsnGlnGlyGly 323
```

RESULT 12

```
US-07-955-905A-23
; Sequence 23, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Gossypium hirsutum
; FEATURE:
```

```
; NAME/KEY: Protein
; LOCATION: 1..587
; OTHER INFORMATION: /note= "Vicilin from G. hirsutum"
US-07-955-905A-23

Alignment Scores:
Pred. No.: 3.57e-06 Length: 587
Score: 156.50 Matches: 112
Percent Similarity: 30.60% Conservative: 78
Best Local Similarity: 18.04% Mismatches: 184
Query Match: 5.82% Indels: 247
DB: 1 Gaps: 23

US-10-728-323-3 (1-1524) x US-07-955-905A-23 (1-587)
Qy 232 TTTGGGTTGATA-----TTTCCTGGT-----TGTCTT 258
Db 19 PheGlyLeuLeuCysSerAlaLysAspPheProGlyArgArgGlyAspAspAspProPro 38
Qy 259 AGACACTATGAA-----GAGCCTCACACACAAAGGTGCTCGATCTCAG 300
Db 39 LysArgTyrGluAspCysArgArgCysGluTTPAspThrArgGlyGlnLysGluGln 58
Qy 301 TCCCAA-----AGACCACCAAGACGTCTCCAA 327
Db 59 GlnGlnCysGluGluSerCysLysSerGlnTyrGlyGluLysAspGlnGlnArgHisArg 78
Qy 328 GGAGAGACCAAGCCCAACAGCAGATAGTACACAGAGGTGCACCGTTTCGATGAG 387
Db 79 ProGluAspProGlnArgArgTyrGluGluCysGlnGlnGluCysArgGlnGlnGluGlu 98
Qy 388 GGTGATCTCATTCGACGTTCCACCGGTGTTGCTTTCTGGCTCTACAAACGACCACGACT 447
Db 98 ----- 98
Qy 448 GATGTTGTTGCTGTTTCTTCTTACTGACACCAACAAACAGCAACACAGCTTGTATGATT 507
Db 99 -----ArgGlnArgProGlnCys 104
Qy 508 CCCAGGAGATTCAATTGCTGGGAACACGAGAGAGTCTTAAAGTTACCAAGCAACAA 567
Db 105 GlnGlnArgCys-----IleLysArgPheGluGlnGln 115
Qy 568 AGCACAAGACGAGACGAAAGACTTTACCATATAGCCCATACAGCCCGCAAGTCTAGCCT 637
Db 116 GlnGlnSerGlnArgGlnPheGlnGluCysGlnGlnHisCysHisGlnGlnGln 135
Qy 628 AGACAAGAGAGCGTGAA-----TTTAGCCCTCGAGGA 660
Db 136 ArgProGluArgLysGlnGlnCysValAlaGluCysArgGluArgTyrGlnGluAsnPro 155
Qy 661 CAGCAGCGCGCAGAGACGAGCAGAGCAAGAGAGAAAGAGGTGGAACATCTTC 720
Db 156 TrpArgArgGluArgGluGluGluAlaGluGluGluThrGluGluGluGluGlu 175
Qy 721 AGCGGCTTCACGCGC---GAGTTCTCTGGAAACAGCCCTCCAGGTTTACACAGACAG--- 774
Db 176 GlnSerHisAsnProPheHisPheHisArgSerPheGlnSerArgPheArgGluGlu 195
Qy 774 ----- 774
Db 196 HisGlyAsnPheArgValLeuGlnArgPheAlaSerArgHisProIleLeuArgGlyIle 215
Qy 774 ----- 774
Db 216 AsnGluPheArgLeuSerIleLeuGluAlaAsnProAsnThrGluValLeuProHisHis 235
Qy 775 -----ATAGTGCAAAACCTAAGAGGC----- 795
Db 236 CysAspAlaGluLysIleTyrLeuValThrAsnGlyArgGlyThrLeuThrPheLeuThr 255
Qy 796 -----GAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACACTGAGGGAGGC 843
Db 796 -----GAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACACTGAGGGAGGC 843
```

256 HisGluAsnLysGluSerTyrAsnValValProGlyValValArgValProAlaGly 275
QY 844 -----CTCAGATCTTGAAGCCAGATAGAAG----- 870
Db 276 SerThrValTyrLeuAlaAsnGlnAsnLysGluLysLeuIleAlaValLeuHis 295
QY 871 AGACGTCCGACGAGGAAGAGGATACGATCA----- 903
Db 296 ArgProValAsnAsnProArgGlnPheGluGluPheProAlaGlySerGlnArgPro 315
QY 903 ----- 903
Db 316 GlnSerTyrLeuArgAlaPheSerArgGluIleLeuGluProAlaPheAsnThrArgSer 335
QY 904 -----GATGAATATGATACGATGAAGAGGATAGAAGCGTGGCAGGGAAGCAGA 954
Db 336 GluGlnLeuAspGluLeuPheGlyGlyArgGlnSerHisArgGlnGlnGlyGlnGly 355
QY 955 GGCAGGGGGAATGCTATTGAAGAGCAGCATCTGCACCGCAAGTGTCTAAAGAACATTGGT 1014
Db 356 MetPheArgLysAlaSerGlnGlnGlnIleArgAlaLeuSerGlnGluAlaThrSerPro 375
QY 1015 AGAACACAGATCCCTGAC-----ATCTACAAC 1041
Db 376 ArgGluLysSerGlyGluArgPheAlaPheAsnLeuLeuTyrArgThrProArgTyrSer 395
QY 1042 CCTCAAGCTGGT-----TCATCAAAATGCCAAGCATCTC 1077
Db 396 AsnGlnAsnGlyArgPheTyrGluAlaCysProArgGluPheArgGlnLeuSerAspIle 415
QY 1078 AACCTTCTAATCTAGTGGCTTGGACCTAGCTGCTGAATATGGAATCTCTACAGGAAT 1137
Db 416 AsnValThrVal-----SerAlaLeu-----GlnLeuAsnGlnGly 427
QY 1138 GCATTGTTTCTGCTCACTACAACACCAAGCAGCAGCATATATATCGATTGAGGGGA 1197
Db 428 SerIlePheValProHisTyrAsnSerLysAlaThrPheValValLeuValAsnGluGly 447
QY 1198 CGGCTCACGTGCAAGTCGTGGAC----- 1221
Db 448 AsnGlyTyrValGluMetValSerProHisLeuProArgGlnSerSerPheGluGluGlu 467
QY 1222 -----AGCAACGGCAACAGATG 1239
Db 468 GluGluGlnGlnGluGlnGluGlnGluGluArgArgSerGlyGlnTyrArgLys 487
QY 1240 TACGACGAGGAGCTTCAAGAGGGTCACGTCTGTGTGTGTCACAGAACTTCGCGCGCT 1299
Db 488 IleArgSerGlnLeuSerArgGlyAspIlePheValValProAlaAsnPheProValThr 507
QY 1300 ---GGAAAGTCCAGACGAGAACTTCGAATACGTGGCATTC----- 1338
Db 508 PheValAlaSerGlnAsnGlnAsnLeuArgMetThrGlyPheGlyLeuTyrAsnGlnAsn 527
QY 1339 ---AAGACAGACTCAAGGCCAGCATAGCAACTCGCGGTGAAACTCCGTCATAGAT 1395
Db 528 IleAsnProAspHisAsnGlnArgIlePhe---ValAlaGlyLysIleAsnHisValArg 546
QY 1396 AACCTGCGGAGGAGGTGTTCGAAATTCATATGCGCTCCAAAGGGAGCAGGCAAGCGAG 1455
Db 547 GlnTrpAspSerGlnAlaLysGluLeuAlaPheGlyValSerSerArgLeuValAspGlu 566
QY 1456 CTTAAGACAAACCCCTTCAAG---TTCTTGTTCACCGTCTCAGCAGTCTCCGAGG 1512
Db 567 IlePheAsnAsnAsnProGlnGluSerTyrPheVal-----SerArgGlnArgGlnArg 584
QY 1513 GCT 1515
Db 585 Ala 585

RESULT 13

US-07-955-905A-2

; Sequence 2, Application US/07955905A

; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-955-905A-2

Alignment Scores:
Pred. No.: 3,91e-06 Length: 566
Score: 156.00 Matches: 106
Percent Similarity: 34.82% Conservative: 74
Best Local Similarity: 20.50% Mismatches: 205
Query Match: 5.80% Indels: 132
DB: 1 Gaps: 22

US-10-728-323-3 (1-1524) x US-07-955-905A-2 (1-566)

QY 253 TGTCTTAGACACTATCAAGAGCCTCACACAGAGTCGTGATCTCATGCCCAAGACCA 312
Db 66 CysGluArgGluTyrLysGluGlnArgGln---GlnGluGluGluGlnArgGln 84
QY 313 CCAAGACGTCTCAAGGAGAAGACCAAGCCACACAGCA----- 351
Db 85 TyrGlnGlnCysGlnGlyArgCysGlnGluGlnGlnGlnGlyGlnArgGluGlnGln 104
QY 352 -----CGAGATAGTCACCAAGAGGTG 372
Db 105 CysGlnArgLysCysTrpGluGlnTyrLysGluGlnArgGlyGluHisGluAsnTyr 124
QY 373 CACCGTTTC-----GATGAGGT-----GATCTCAT 399
Db 125 HisAsnHisLysAsnArgSerGluGluGluGlyGlnGlnArgAsnAsnProTyr 144
QY 400 GCAGTTCACCGGTGTTCTTCTGGCTC---TACAACGACCAACACTGATGTGTT 456
Db 145 TyrPheProLysArgArgSerPheGlnThrArgPheArgAspGluGlyAsnPheLys 164
QY 457 GCTGTTCTTCTTACTCACACCAACAC-----NACCAACACAGCTT 498
Db 165 IleLeuGlnArgPheAlaGluAsnSerProProLeuLysGlyIleAsnAspTyrArgLeu 184
QY 499 GATCAGTTC-----CCGAGGATTCATTTGCTGGGAACACGAGCAAGAG--- 546
Db 185 AlaMetPheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAla 204
QY 547 -----TTCTTAGGTACCAACCAAGC 570
Db 205 IleTyrPheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGlu 224
QY 571 AGCAAAAGCAGACGAGA-----AGCTTACCATAT---AGCCATACACCCCG 615
Db 225 SerTyrAsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrVal 244
QY 616 CAAAGTCAGCTAGACAAGAGAGCGTGAATTAGC-----CCTCAGGACAG 663
Db 245 ValSerGlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSer 264

```
QY 664 CACAGCCGACAGAACGAGCAGCAGAACAGAAAGAAACGAGGTCGAAACATCTTTCAGC 723
Db   :|||
QY 265 ProGlyLysTyrGluLeuPheProAlaGlyAsnAsnLysProGluSerTyrTyrGly 284
Db   :|||
QY 724 GCGTTCACGCCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGAC-----AGACAG 774
Db   :|||
QY 285 AlaPheSerTyrGluValLeuGluThrValPheAsnThrGlnArgGluLysLeuGluGlu 304
Db   :|||
QY 775 ATAGTCAAAACCTAGAGCCGAGACCGGAGAGTGAAGAGGAGGAGGAGCCATCTGTGACAGTG 834
Db   :|||
QY 305 IleLeuGluGluGlnArgGlyGlnLysArgGlnGlnGlyGln----- 318
QY 835 AGGGGAGCGCTCAGATCTTTGAGCCCGACAGATAGAAAGAGAGCTGCCGACGAAAGAGAGAA 894
Db   :|||
QY 319 GlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIleSerGlnGlnAla--- 337
QY 895 TACGATGAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 954
Db   :|||
QY 338 -----ThrSerProArgHisArgGly---GlyGluArg 347
QY 955 GGCAGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAACATTTGGT 1014
Db   :|||
QY 348 LeuAlaIleAsnLeuLeu----- 353
QY 1015 AGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACTCAAAACTGCCAACAGAT 1074
Db   :|||
QY 354 ---SerGlnSerPro---ValTyrSerAsnGlnAsnGlyArgPhePheGluAlaCysPro 371
QY 1075 CTCACCTCTTAATACTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1134
Db   :|||
QY 372 GluAspPheSerGlnPheGlnAsnMetAspValAlaValSerAlaPheLysLeuAsnGln 391
QY 1135 AATGCAATGTTTTCGCTGCTACTACACCAACCAACGACGACATCATATATGATTGAGG 1194
Db   :|||
QY 392 GlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValValPheValThrAsp 411
QY 1195 GGACGGGCTCACGTGCAAGTCGTG-----CACAGCAACGGCAAC--- 1233
Db   :|||
QY 412 GlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSerGlnGlySerGln 431
QY 1234 -----AGATGTGACGAGGAGGCTTCAAG----- 1260
Db   :|||
QY 432 SerGlyArgGlnAspArgArgGluGlnGluSerGluGluSerGluGluThrPheGlyGlu 451
QY 1261 -----GGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
Db   :|||
QY 452 PheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAlaProAlaGlyHis 471
QY 1291 GCGGTC-----GCTGGAAGTCCCAGAGCGAGAATTCGAATACGTGGCATTTCAAG 1341
Db   :|||
QY 472 AlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAlaPheGlyLeuAsn 491
QY 1342 ACAGACTCAAGGCCGACATAGCAACCTCGCGGTGAAACCTCCGTCATAGATAACCTG 1401
Db   :|||
QY 492 AlaGlnAsnAsnGlnArgIlePhe---LeuAlaGlyLysLysAsnLeuValArgGlnMet 510
QY 1402 CCGGAGGAGGTGTGCAAAATTCATATGGCTCCAAAGGAGGAGCAGCAAGCGAGCTTAAG 1461
Db   :|||
QY 511 AspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuValAspAsnIlePhe 530
QY 1462 AACAAACACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGTCTCCGAGG 1512
Db   :|||
QY 531 AsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArgGlnArg 547
```

RESULT 14

```
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: theobroma cacao
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..566
; OTHER INFORMATION: /note="67 kD Precursor Protein"
; US-07-955-905A-22
```

Alignment Scores:

Pred. No.:	3,91e-06	Length:	566
Score:	156.00	Matches:	106
Percent Similarity:	34.82%	Conservative:	74
Best Local Similarity:	20.50%	Mismatches:	205
Query Match:	5.80%	Indels:	132
DB:	1	Gaps:	22

US-10-728-323-3 (1-1524) x US-07-955-905A-22 (1-566)

```
QY 253 TGTCTAGACATATCAAGAGCCTCACACACAGGTGTCGATCTCAGTCCCAAGACCA 312
Db   :|||
QY 66 CysGluArgGluTyrLysGluGlnArgGln---GlnGluGluGluGlnArgGln 84
Db   :|||
QY 313 CCAAGACGTCTCCAAGGAGAAGACCAAGCCCAACAGCAA----- 351
Db   :|||
QY 85 TyrGlnGlnCysGlnGlyArgCysGlnGluGlnGlnGlnGlnArgGlnGlnGln 104
Db   :|||
QY 352 -----CGAGATAGTCACCAAGAGGTG 372
Db   :|||
QY 105 CysGlnArgLysCysTrpGluGlnTyrLysGluGlnArgGlyGluHisGluAsnTyr 124
Db   :|||
QY 373 CACCGTTTC-----GATGAGGT-----GATCTCAT 399
Db   :|||
QY 125 HisAsnHisLysLysAsnArgSerGluGluGluGluGlyGlnGlnArgAsnAsnProTyr 144
QY 400 GCAGTTCCCAACCGGTGTTCTTTCTGGCTC---TACACGACGACGACACTGATGTTGT 456
Db   :|||
QY 145 TyrPheProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPheLys 164
QY 457 GCTGTTTCTTACTGACACCAAC-----AACGACAAACAGCTT 498
Db   :|||
QY 165 IleLeuGlnArgPheAlaGluAsnSerProProLeuLysGlyIleAsnAspTyrArgLeu 184
QY 499 GATCAGTTC-----CCAGGAGATTCAATTTGGCTGGGAACACGAGCAAGAG--- 546
Db   :|||
QY 185 AlaMetPheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAla 204
QY 547 -----TTCTTAAGGTACCAACCAAAAGC 570
Db   :|||
QY 205 IleTyrPheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGlu 224
QY 571 AGACAAGACGACGACGA-----ACGTACCATAT---AGCCATACAGCCCG 615
Db   :|||
QY 225 SerTyrAsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrVal 244
QY 616 CAAAGTCACCTAGACAAGAAGCGGTGAATTTAGC-----CCTCGAGACAG 663
Db   :|||
QY 245 ValSerGlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSer 264
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 24, 2005, 10:05:17 ; Search time 98.2593 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 cgcagcaaccggaggagaa.....ctccgaggctgtggtcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124357_29323/app_query.fasta_1.4757
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10728323@cgn_1_1_221@runat_23082005_124357_29323
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubppa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubppa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2673	99.3	507	15	US-10-245-871-96	Sequence 96, Appl
2	2673	99.3	507	15	US-10-253-286-96	Sequence 96, Appl
3	2673	99.3	510	14	US-10-228-806-6	Sequence 6, Appl
4	2673	99.3	510	15	US-10-100-303A-90	Sequence 90, Appl
5	2638	98.0	526	9	US-09-731-221-79	Sequence 79, Appl
6	2621	97.4	530	17	US-10-899-551-6	Sequence 6, Appl
7	1665	61.9	351	17	US-10-899-551-58	Sequence 58, Appl
8	1502	55.8	481	10	US-09-759-967-21	Sequence 21, Appl
9	1502	55.8	481	15	US-10-424-599-171702	Sequence 171702, A
10	1502	55.8	488	15	US-10-425-114-43865	Sequence 43865, A
11	1502	55.8	491	15	US-10-425-114-44047	Sequence 44047, A
12	1502	55.8	492	15	US-10-425-114-43836	Sequence 43836, A
13	1502	55.8	492	15	US-10-425-114-46294	Sequence 46294, A
14	1499.5	55.7	466	15	US-10-425-114-71959	Sequence 71959, A
15	1499.5	55.7	466	15	US-10-425-114-71973	Sequence 71973, A
16	1499.5	55.7	485	10	US-09-759-967-20	Sequence 20, Appl
17	1499.5	55.7	488	15	US-10-425-114-45645	Sequence 45645, A
18	1499.5	55.7	489	15	US-10-425-114-43983	Sequence 43983, A
19	1499.5	55.7	489	15	US-10-425-114-44676	Sequence 44676, A
20	1499.5	55.7	489	15	US-10-425-114-45612	Sequence 45612, A
21	1499.5	55.7	489	15	US-10-425-114-50117	Sequence 50117, A
22	1499.5	55.7	489	15	US-10-425-114-50207	Sequence 50207, A
23	1499.5	55.7	489	15	US-10-425-114-51412	Sequence 51412, A
24	1499.5	55.7	489	15	US-10-425-114-58429	Sequence 58429, A
25	1499.5	55.7	489	15	US-10-425-114-71902	Sequence 71902, A
26	1499.5	55.7	489	15	US-10-425-114-71909	Sequence 71909, A
27	1499.5	55.7	489	15	US-10-425-114-71957	Sequence 71957, A
28	1499.5	55.7	489	15	US-10-425-114-71958	Sequence 71958, A
29	1499.5	55.7	489	15	US-10-425-114-71967	Sequence 71967, A
30	1499.5	55.7	489	15	US-10-425-114-71995	Sequence 71995, A
31	1499.5	55.7	489	15	US-10-425-114-72287	Sequence 72287, A
32	1499.5	55.7	489	15	US-10-425-114-72288	Sequence 72288, A
33	1498	55.6	476	17	US-10-425-114-45639	Sequence 45639, A
34	1496	55.6	477	17	US-10-409-993-1	Sequence 1, Appl
35	1496	55.6	477	17	US-10-409-993-15	Sequence 15, Appl
36	1496	55.6	485	17	US-10-409-993-9	Sequence 9, Appl
37	1496	55.6	485	17	US-10-409-993-13	Sequence 13, Appl
38	1496	55.6	495	15	US-09-759-967-2	Sequence 2, Appl
39	1496	55.6	495	15	US-10-424-599-171705	Sequence 171705, A
40	1496	55.6	495	15	US-10-425-114-45626	Sequence 45626, A
41	1496	55.6	501	15	US-10-425-114-43904	Sequence 43904, A
42	1496	55.6	505	15	US-10-425-114-71930	Sequence 71930, A
43	1496	55.6	505	15	US-10-425-114-71962	Sequence 71962, A
44	1496	55.6	506	15	US-10-425-114-43901	Sequence 43901, A
45	1496	55.6	506	15	US-10-425-114-45624	Sequence 45624, A

ALIGNMENTS

RESULT 1
US-10-245-871-96
; Sequence 96, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT

ORGANISM: Arachis hypogaea
US-10-245-871-96

Alignment Scores:

Pred. No.: 1,83e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-3 (1-1524) x US-10-245-871-96 (1-507)

QY	1	CGGCAGCAACGGAGGAGACGGCTGCCAGTTCACGGCGCTCAATGCGCAGAGACCTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCATGAATCAGAGGCGGTTCATTAGACTTGGAAACCCCAACACGAGGATTC	120
DB	21	AsnArgileGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCTCGCGGCAACGCCCTTCGTAGGCCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTTG	240
DB	61	PheTy-SerAsnAlaProGlnGluillePheilleGlnGlnGlyArgGlyTyPheGlyLeu	80
QY	241	ATATTCCTGGTTGCTTAGCACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG	300
DB	81	IlePheProGlyCysProArgHisTyrgluGluProHisThrGlnGlyArgArgSerGln	100
QY	301	TCCCAAGACCACAGACGTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT	360
DB	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCACCGCGGTGTGCT	420
DB	121	HisGlnIysValHisArgPheAspGluGlyAspLeulleAlaValProThrGlyValAla	140
QY	421	TTCTGGCTTACACGACCAACGACACTGATGTTGTTGCTGTTTCTTCTTACTGACACAA	480
DB	141	PheTrpLeuTyrrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
QY	481	AACAACGACACCGCTTGATCAGTTCCTCCAGGAGATTCATTTGCTGGGACACGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACGAGCAAGAGAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrglnGlnSerArgGlnSerArgArgSerLeuProTyrr	200
QY	601	AGCCCATACAGCCGCAAGTACGCTAGCAAGAGAGCGTGAATTTAGCCCTCAGGA	660
DB	201	SerProTyrrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyAsnIlePhe	240
QY	721	AGCGGCTTACGCGCGGAGTTCTTGGAAACAGCCTTCAGGTTTCACGACACAGATAGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAAACCTAAGAGCGGAGCCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280
QY	841	GGCCTCAGATCTTGACCCAGATAGAAAGAGACGTGCGACCAAGAGGAATACGAT	900
DB	281	GlyLeuArgileLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrrAsp	300

QY	901	GAAGATGAATATGAATACGATGAAGAGGATAGACGGCTGGCAGGGGAGAGAGCGCAGG	960
DB	301	GluAspGluTyrrGluTyrrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAGAACATTGGTGAAC	1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
QY	1021	AGATCCCTCGACATCTACAACCCCTCAAGCTGTTTCACTCAAACTGCCAACCATCTCAAC	1080
DB	341	ArgSerProAspIleTyrrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAATACTTAGTGGCTTCGACCTAGTGTGAATATGGAAATCTCTACAGGAATGCA	1140
DB	361	LeuLeulleLeuArgTyrrLeuGlyProSerAlaGluTyrrGlyAsnLeuTyrrArgAsnAla	380
QY	1141	TTGTTTTCGCTCACTACAACCAACGACACACATCATATATCGATTGAGGGGACGG	1200
DB	381	LeuPheValAlaHisTyrrAsnThrAsnAlaHisSerIleIleTyrrArgLeuArgGlyArg	400
QY	1201	GCTCAGCTGCAAGTCTGGACAGCAACGCAACAGAGTGTACGACGAGGCTTCAAGAG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrrAspGluGluLeuGlnGlu	420
QY	1261	GCTCAGCTGCTTGTGTCACAGAACTTCGCCGCTCGCTGGAAAGTCCAGAGCGAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTGGAATAGCTGGCATTCAAGACACTCAAGGCCACGATAGCCAACTCCCGCGTGA	1380
DB	441	PheGluTyrrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCCGTCATAGATAAACCCTCGCGAGGAGGTGTTGCAAAATTCATATGGCCTCCAAAGG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrrGlyLeuGlnArg	480
QY	1441	GAGCAGGCAAGCAGCTTAAAGAACAAACCCCTTCAAGTTCCTCCACCGTCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	501	GlnSerProArgAlaValAla	507

RESULT 2

US-10-253-286-96
; Sequence 96, Application US/10253286
; Publication No. US20040058861A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-96

Alignment Scores:
Pred. No.: 1.83e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0

DB: 15 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-253-286-96 (1-507)
QY 1 CGGCAGCAACCGGAGGAGAACCGCTGCGAGTTCCAGCGCCCTCAATGCGCAGACCTGAC 60
DB 1 ArgGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY 61 AATCGCATTTGAATCAGAGGGCGGTACATTAGACTTGGAAACCCCAACACCGAGATTC 120
DB 21 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 40
QY 121 GAATCGCGCGCGCTCTCTCGCTTGTAGTCTCCGCGCAACCGCCCTCGTAGGCCT 180
DB 41 GluCysAlaGlyAlaAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 60
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGGTTG 240
DB 61 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY 241 ATATTCCCTGTTGCTTAGACACTATGAAGAGCCTCACACAAAGTGTGTCGATCTCAG 300
DB 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 100
QY 301 TCCCAAGACACCAAGACGCTCTCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
DB 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY 361 CACCAGAGGTGACCGCTTTCGATGAGGTGATCTCATTTGCAGTTCCACCGGTTGCT 420
DB 121 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY 421 TTCTGGCTCTACAGCACCGACGACACTGATGTTGCTGCTGTTCTTACTGACACCAAC 480
DB 141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY 481 AACCAACGACACACGCTGATCAGTCCCGAGGATTCATTTGGCTGGGAACACCGAG 540
DB 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY 541 CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
DB 181 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 200
QY 601 AGCCCATACAGCCCGCAAGTACGCTAGACAAGAGCGGTGAATTTAGCCCTCGAGGA 660
DB 201 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 220
QY 661 CAGCACAGCCGACAGACGAGCAGGACAGCAAGAAACGAGGTGGAACATCTTC 720
DB 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 240
QY 721 AGCGGCTTCCGCGGAGTTCCTCGAAACAGCCCTCCAGGTTGACGACAGACAGATAGTG 780
DB 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY 781 CAAACCTTAAGACGAGACCGAGAGTGAAGAAGAGGAGCCATTGTACAGTCAGGGGA 840
DB 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 280
QY 841 GGCTTCAGATCTTTAGCCCGCCAGATAGAAAGAGAGCTGCCAGCAGAGAGGATACGAT 900
DB 281 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 300
QY 901 GAAGATGAATATGAATACGATGAAGAGATAGAGGCTGCGAGGGGAAGCAGAGGAGG 960
DB 301 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 320
QY 961 GGGATCGTATTGAAGAGACGATCTGACCCCGAGTCTAAAAAGAAACATTGGTAGAAC 1020
DB 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340
QY 1021 AGATCCCTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACTGCCAACGATCTCAAC 1080

DB 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY 1081 CTTCTAATATCTAGTGGCTTGACCTAGTCTGAATATGAAATCTCTACAGAAATGCA 1140
DB 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380
QY 1141 TTGTTGTGCTGCTCACTAACACCAACGACGACAGCATCATATATCGATTGAGGGGACGG 1200
DB 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY 1201 GCTCAGTGAAGTCTGTCGACCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
DB 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
QY 1261 GGTCACTGCTGCTGTCGTCACAGAACTTCCGCGTGCCTGGAAGTCCACAGAGCGAAC 1320
DB 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY 1321 TTCGAATACGTGGCATTCAGACAGACTCAAGGCCGACGATAGCCAACTTCGCCGGTGAA 1380
DB 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY 1381 AACTCCGTCTATAGTAACTGCGGAGGAGTGTGTCGAAATTCATATGCGCTCCAAAGG 1440
DB 461 AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg 480
QY 1441 GACGAGCGAAGCAGCTTAAGAAACAACACCCCTTCAAGTCTTCCGTTCCACCGTCTCAG 1500
DB 481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
DB 501 GlnSerProArgAlaValAla 507
RESULT 3
US-10-228-806-6
; Sequence 6, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-6
Alignment Scores:
Pred. No.: 1,83e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 14 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-228-806-6 (1-510)
QY 1 CGGCAGCAACCGGAGGAGAACCGTCCAGTTCACGCGCTCAATGCGCAGACCTGAC 60
DB 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGATTC 120
DB 24 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATCGCGCGGCTGTCGCCCTCTCTCGCTTAGTCTCCGCGCAACCGCCCTTCGTAGGCCT 180

Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCCAAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACACAAGTCTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCCAAAGACTCTCCAAGGAGAGACAAAGCCAAAGCCAAAGCAGATAGT 360
Db 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACACAGCACACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480
Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACAAAGACACCAAGCTTGATCAAGTTCCCGAGAGATTCAAATTTGGCTGGAAACAGGAG 540
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCTTAAGTACCAGCAACAAGCAGCAACAAGCAGCAACAAGCTTACCATAT 600
Db 184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY 601 AGCCATACAGCCCGCAAGTCTAGCTTAGACAAGAGAGCGTCAATTTAGCCCTCGAGGA 660
Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223
QY 661 CAGCACAGCCGAGAGAACAGCAGGAGCAAGAAGAAAGAAACGAAGTGGAAACATCTTC 720
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
QY 721 AGCGCTTACCGCGGNGTCTCGACAGCCCTCCAGGTTCACGACAGACAGATAGTG 780
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAAACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGGGGA 840
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 283
QY 841 GGCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTGCCGACCGAAGAGGAATACGAT 900
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
QY 901 GAAGATCAATATGAATACGATGAAGGATAGAGCGGTGGCAGGAGGAGCAGGAGG 960
Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 323
QY 961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAAAGAACATTGTAGAAAC 1020
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCATCTCAAACTGCCAACGATCTCAAC 1080
Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATACTAGTGGCTTGACCTAGTGTGAATATGAAATCTCTACAGGAATGCA 1140
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
QY 1141 TTGTTGTCTCCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY 1201 GCTCAGTGAAGTCTGTGACAGCAAGCGGCAACAGAGTGTACACAGAGGCTTCAAGAG 1260

Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY 1261 GGTCAAGTCTGTCGTGTCACAGAACTTCGCGCTCGTGGAAAGTCCCGAGAGCGAAGC 1320
Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
Db 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCGTATAGATAAACCCTCCGAGAGGTGGTTCGAAATTCATATGCGCTCCAAAGG 1440
Db 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
QY 1441 GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
Db 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 504 GlnSerProArgAlaValAla 510
RESULT 4
US-10-100-303A-90
; Sequence 90, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-90
Alignment Scores:
Pred. No.: 1.83e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 15 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-100-303A-90 (1-510)
QY 1 CGGACGACACCGAGAGAACCGTGCAGTTCAGCGCTCAATCGCGAGAGACCTGAC 60
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTAATACAGAGCGGTTCATTGAGACTTGGAAACCCCAACACAGAGATTC 120
Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGGCGTGCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTGTAGGCCT 180
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCCAAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACACAAGTCTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
QY 301 TCCCAAGACCCAAAGACTCTCCAAGGAGAGACAAAGCCAAAGCCAAAGCAGATAGT 360

Db	104	SerGlnArgProProArgArgLeuGlnGlyAspGlnSerGlnGlnGlnArgAspSer	121
Qy	361	CACCAGAAAGGTGCACCGTTTCATCAGGGTGATCTCATTGCAGTTGCCACCGGTGCTGCT	420
Db	124	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
Qy	421	TTCTGGCTCTCAACGACACGACACTGATGTGTGTGCTGTTTCTCTTACTGACACCAAC	480
Db	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
Qy	481	AACAAACGACACACCGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAAACACGAG	540
Db	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
Qy	541	CAAGAGTTCTTAAAGTACACGACAAACAAAGCAGACAAAGCAGCAAGAAAGCTTACCATAT	600
Db	184	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	203
Qy	601	AGCCCATACGCCCGCAAAAGTCAGCCTAGACAAGAAGACGCTGAATTTAGCCCTCGAGGA	666
Db	204	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	223
Qy	661	CAGCACGCCCGCAGACGACGACGACCAAGACAGAGAAACCAAGAGTGGAAACATCTTC	720
Db	224	GlnHisSerArgArgLeuArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe	243
Qy	721	AGCCGCTTCACGCCGAGTCTCTGCAACAGCCTTCAGGTTGACGACAGACAGACATAGTG	780
Db	244	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	263
Qy	781	CAAAACCTTAAGAGCGAGACCCGACAGATGAAGAAGAGGAGGCCATTGTGCACAGTGGGGA	840
Db	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly	283
Qy	841	GGCCTCAGAACTTCGACCCAGATAGAAAGAGACGTCGCCGACGAAGAAGAGCAATACGAT	900
Db	284	GlyLeuArgIleLeuSerProAspArgIysArgArgAlaAspGluGluGluGluTyrAsp	303
Qy	901	GAAGATGAATATGAATACGATTAAGAGGATAGAAGCGTCGCAGGGAAAGCAGAGGCAGG	960
Db	304	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	323
Qy	961	GGGAATGGTATTGAAGAGACGATCTGCACCCGCAAGTCTAAAAAGAACATTTGGTAGAAAC	1020
Db	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	343
Qy	1021	AGATCCCTTGACATCTCAACCCCTCAGCTGGTTCACTCAAAAAGTCCCAACGATCTCAAC	1080
Db	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuTyrThrAlaAsnAspLeuAsn	363
Qy	1081	CTTCTAATACTTAGTGCTGGACCTAGTCTGAATGAAATCTCTACAGGAATGCA	1140
Db	364	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
Qy	1141	TTGTTTGTGCTCACTCAACACCAACGACACACGACATCATATATCGATTGAGGGGACGG	1200
Db	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
Qy	1201	GCTCAGCTGCAGTCTGGACAGCAACGGCAACAGAGTGTTACGACGAGGAGCTTCAAGAG	1260
Db	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	423
Qy	1261	GGTCACGTGCTTGTGTGCCACAGAACTTCGCCGTGCGTGGAAAGTCCCAGAGCGAGAAC	1320
Db	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
Qy	1321	TTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGACATAGCAACCTCGCGGTCAA	1380
Db	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
Qy	1381	AACTCCGCTCATAGATAACTGCCGAGAGGTGTTGGCAATTCATATGGCCTCAAGG	1440
Db	464	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	483

QY	1441	GAGCAGGCAAGCGAGCTTAAGAAACAACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG	1500
Db	484	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln	503
QY	1501	CAGTCTCCGAGGGCTGTGGCT	1521
Db	504	GlnSerProArgAlaValAla	510
RESULT 5			
US-09-731-221-79			
; Sequence 79, Application US/09731221			
; Patent No.. US20020018778A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael			
; TITLE OF INVENTION: Passive Desensitization			
; FILE REFERENCE: 2002834-0103			
; CURRENT APPLICATION NUMBER: US/09/731,221			
; CURRENT FILING DATE: 2001-12-06			
; NUMBER OF SEQ ID NOS: 79			
; SOFTWARE: Patent In Ver. 2.1			
; SEQ ID NO 79			
; LENGTH: 526			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence:Arachis			
; OTHER INFORMATION: Hypogaea			
US-09-731-221-79			
Alignment Scores:			
Pred. No.: 1,94e-218 Length: 526			
Score: 2638.00 Matches: 502			
Percent Similarity: 99.01% Conservative: 0			
Best Local Similarity: 99.01% Mismatches: 5			
Query Match: 98.03% Indels: 0			
DB: 9 Gaps: 0			
US-10-728-323-3 (1-1524) x US-09-731-221-79 (1-526)			
QY	1	CGGCAGCAACGGAGGAGAACGGGTGCCAGTCTCCAGCGCTCAATGCGCAGACCTGAC	60
Db	5	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	24
QY	61	AATCGCATTAATCAGAGGGCGGTACATTGACACTTGGAACCCCAACACACGAGGAGTTC	120
Db	25	AsnArgileGluSerGluGlyGlyTyrlleGluThrTrpAsnAlaAsnGlnGluPhe	44
QY	121	GAATGCCCGCGCGCGCTCTCGGTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCCT	180
Db	45	GluCysAlaGlyValAlaLeuSerArgLeuValleuArgAsnAlaLeuArgPro	64
QY	181	TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTG	240
Db	65	PheTySerAsnAlaProGlnGluilePheileGlnGlnGlyArgGlyTyrPheGlyLeu	84
QY	241	ATATTCTCTGGTGTCTCTAGACATATGAGAGCGCTCACACAAAGTCGTGCATCTCAG	300
Db	85	IlePheProGlyCysProArgHisTyrlleGluGluProHisThrGlnGlyArgArgSerGln	104
QY	301	TCCCAAGACACCAAGACGCTCTCCAAGGAGAAGACCAAGCCCAACAGCAACGAGATAGT	360
Db	105	SerGlnArgProProArgArgLeuGlnGlyValaspGlnSerGlnGlnArgAspSer	124
QY	361	CACCAGAAGTGACCGTTTCGATGAGGGTGATCTCATTCAGCTTCCCACCGGTGTGCT	420
Db	125	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	144
QY	421	TTCTCGCTCTACAACGACACGACATGATGTTGTTGCTTCTCTTACTACACCAAC	480
Db	145	PheTrpLeuTyrlleAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	164
QY	481	AACAACGACAAACAGCTTGATCAGTTTCCCCAGGAGATTCAATTTGGCTCGGAACCGGAG	540

Db	165	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAenThrGlu	184
QY	541	CAAGAGTTCTTAAGTACCAAGCAACAAAGCAGACGAAGAAGCTTACCATAT	600
Db	185	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	204
QY	601	AGCCCATACAGCCCGCAAGTCAGCCCTAGACAAGAAGCGTGAAATTTAGCCCTCAGGA	660
Db	205	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	224
QY	661	CAGCAGCGCCAGAGAACAGCAGACGAAGAAGAAAGAAACGAAGGTGGAACATCTTC	720
Db	225	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAenIlePhe	244
QY	721	AGCGGCTTCACCGCGAGTTCCTGGAAACAAGCCCTCCAGGTTGACACACAGACATAGNG	780
Db	245	SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal	264
QY	781	CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
Db	265	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	284
QY	841	GGCCTCAGAACTTGTAGCCAGATAGAAGAGACGTGCCGACGAAGAAGGAATACGAT	900
Db	285	GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	304
QY	901	GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGCTGGCAGGGGAACGAGGCAGG	960
Db	305	GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	324
QY	961	GGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAGAACATTTGGTAGAAAC	1020
Db	325	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleLysAsnIleGlyArgAsn	344
QY	1021	AGATCCCTCGACATCTACAACCTCAAGCTGGTTCTACTCAAACTGCCAACGATCTCAAC	1080
Db	345	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	364
QY	1081	CTTCTAATACTTAAAGTGGCTTGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA	1140
Db	365	LeuLeuIleLeuArgTyrLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	384
QY	1141	TTGTTTTCGCTCACTACAACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
Db	385	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	404
QY	1201	GCTCAGTCGAAGTCGTGGACGACCAACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG	1260
Db	405	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	424
QY	1261	GGTCAGCTGCTTGTGTGCCACAGAATTCGCCGTCGCTGGAAGTCCCGAGCGAGAAC	1320
Db	425	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	444
QY	1321	TTCSAATACGTGGCATTCAAGACAGACTCAAGGCCCGCAGCATAGCCAACTCTCGCGGTGAA	1380
Db	445	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	464
QY	1381	AATCCGTCATAGATAACCTGCCGAGGAGGTGGTTGCCAAATTCATATATGGCTCCAAAGG	1440
Db	465	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	484
QY	1441	GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAG	1500
Db	485	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln	504
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
Db	505	GlnSerProArgAlaValAla	511
RESULT 6			
US-10-899-551-6			

; Sequence 6, Application US/10899551			
; Publication NO. US20050063994A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael J.			
; APPLICANT: Burks, A. Wesley			
; APPLICANT: Sampson, Hugh A.			
; APPLICANT: Howard, Sosin B.			
; APPLICANT: Bottomly, Kim H.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy			
; FILE REFERENCE: 2002834-0233			
; CURRENT APPLICATION NUMBER: US/10/899,551			
; CURRENT FILING DATE: 2004-07-26			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 6			
; LENGTH: 530			
; TYPE: PRT			
; ORGANISM: species Arachis hypogea			
US-10-899-551-6			
Alignment Scores:			
Pred. No.:	5,7e-217	Length:	530
Score:	2621.00	Matches:	500
Percent Similarity:	98.62%	Conservative:	7
Best Local Similarity:	98.62%	Mismatches:	0
Query Match:	97.40%	Indels:	0
DB:	17	Gaps:	0
US-10-728-323-3 (1-1524) x US-10-899-551-6 (1-530)			
QY	1	CGCGACGACACCGAGAGAACCGCTGCCAGCGCTCAATCGCGCAGAGACCTGAC	60
Db	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	AATCCGATTGATCAGAGCGGCTTACATTGAGACTTGGAACCCCAACACAGGAGTTC	120
Db	44	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCCGCAACGCCCTTCGTAGGCCT	180
Db	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	83
QY	181	TTCTACTCTCAATGCTCCCGAGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTG	240
Db	84	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	103
QY	241	ATATTCCCTGTTGCTTAGACACTATGAAGAGCCTCACACACAGGTGCTCGATCTCAG	300
Db	104	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln	123
QY	301	TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGACCAAGCCCAACAGCAGATAGT	360
Db	124	SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer	143
QY	361	CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTCGATGTTCCCAACCGGTGTCT	420
Db	144	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
QY	421	TTCTCGCTCTACAAGCCACGACCTGATGTTGTTGTTGTTTCTTCTTACTCACACCAAC	480
Db	164	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	183
QY	481	AACAACGACCAACCCAGCTGATCAGTTCCCGAGGAGATTCAATTTGGCTGGGAACCGGAG	540
Db	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
QY	541	CAAGAGTCTTAAAGGTACAGCAACAAAGCAGACAAAAGCAGACGAAGAGCTTACCATAT	600
Db	204	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	223
QY	601	AGCCCATACAGCCCGCAAGTCCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA	660
Db	224	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	243


```
QY 661 CAGCACAGCCGACAGAACGAGGACGAGACAGAAACGAGAGTGAACATCTTC 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGluArgGlnGluArgGlyGlyAsnIlePhe 263
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAACAAGCCCTCCAGGTTGACGACAGACAGATAGTG 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
Db 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGlyAlaIleValThrValArgGly 303
QY 841 GGCTCAGATCTTGACCCAGATAGAAAGAGAGCGTCCGACGAGAGAGGATACGAT 900
Db 304 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluIuArgAsp 323
QY 901 GAAGATGAATATGAATACGATGAAGAGCATAGAGCGGTGCGAGGGAAGACAGAGGCGG 960
Db 324 GluAspGluTyrgluTyrgluTyrgluAspArgArgArgGlyArgGlySerArgGlyArg 343
QY 961 GGAATCGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAACATTTGGTAGAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 363
QY 1021 AGATCCCTGCATCTACACACCTCAAGCTGGTTCACCTCAAACTGCCAAGCATCTCAAC 1080
Db 364 ArgSerProAspIleTyrgluAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 383
QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTCTGAATATGAAATCTCTACAGGAATGCA 1140
Db 384 LeuLeuIleuArgIlePheGlyProSerAlaGluTyrglyAsnLeuTyrgArgAsnAla 403
QY 1141 TTGTTTTCGCTCACTACACACCAACGACGACGATCATATATCGATTGAGGGGACGG 1200
Db 404 LeuPheValAlaIleTyrgluAsnAlaHisSerIleIleTyrgluArgLeuArgGlyArg 423
QY 1201 GCTCAGTGAAGTCTGGACGACGACGACGACGACGACGACGACGACGACGACGACG 1260
Db 424 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrgluGluGluLeuGlnGlu 443
QY 1261 GGTCACTGCTGTGGTGCCACAGAACTTCGCGTCTGGAAGTCCACGAGCGAGAAC 1320
Db 444 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 463
QY 1321 TTCGAATACGTGGCATTCAAGACAGACTCAAGGCCACGATAGCCTCGCGGTGAA 1380
Db 464 PheGluTyrgluValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
QY 1381 AACTCGTCTAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 484 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrglyLeuGlnArg 503
QY 1441 GAGCAGCAAGGAGGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db 504 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 523
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 524 GlnSerProArgAlaValAla 530
```

RESULT 7

```
US-10-899-551-58
; Sequence 58, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
```

```
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 351
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-58
```

Alignment Scores:

Pred. No.:	1,53e-134	Length:	351
Score:	1665.00	Matches:	315
Percent Similarity:	97.83%	Conservative:	0
Best Local Similarity:	97.83%	Mismatches:	7
Query Match:	61.87%	Indels:	0
DB:	17	Gaps:	0

US-10-728-323-3 (1-1524) x US-10-899-551-58 (1-351)

```
QY 1 CGGCACAAACCGGAGGAGAAACGCGTCCAGTTCACGCGCTCAATGCGCAGAGACTTGAC 60
Db 15 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 34
QY 61 AATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACACGAGGATTC 120
Db 35 AsnArgIleGluSerGluGlyGlyIleGluThrTrpAsnProAsnAsnGlnGluPhe 54
QY 121 GAATCGCGCGGCTCCCGCTCTCTCGCTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 55 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 74
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAAAGGGGATATTTGGGTTG 240
Db 75 PheTySerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrgPheGlyLeu 94
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAGAGCCTCACACAAAGTCTGTCATCTCAG 300
Db 95 IlePheProGlyCysProSerThrTyrgluGluProAlaGlnGlnGlyArgArgSerGln 114
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 115 SerGlnArgProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 134
QY 361 CACCAGAAGGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCCTCCACCGTGTGCT 420
Db 135 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 154
QY 421 TTCTGCTCTACACGACCAACGACACTGATGTGTGTGCTGTTCTTCTTACTGACACCAAC 480
Db 155 PheTrpLeuTyrgAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 174
QY 481 AACCAACGACCAACCGAGTTCATGATCTCCCGAGGAGATTCAATTTGCTGGGAACACGAG 540
Db 175 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 194
QY 541 CAAGAGTCTTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 195 GlnGluPheLeuArgTyrgGlnGlnSerArgGlnSerArgArgSerLeuProTyrg 214
QY 601 AGCCCATACAGCCCGCAAGTCCAGCTCAGCAAGAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 215 SerProTyrgSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 234
QY 661 CAGCAGCGCGCAGAGAACGAGCAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 235 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluPhe 254
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCACAGTTTGACGACAGACAGATAGTG 780
Db 255 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 274
QY 781 CAAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
```


RESULT 9
 US-10-424-599-171702
 ; Sequence 171702, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Cao Yongwei
 ; APPLICANT: Zhou Yihua
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 171702
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_126061C.1.pap
 US-10-424-599-171702

Alignment Scores:
 Pred. No.: 1.99e-120 Length: 481
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.57% Conservative: 68
 Best Local Similarity: 58.24% Mismatches: 91
 Query Match: 55.82% Indels: 54
 DB: 15 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-424-599-171702 (1-481)

QY	1	CGGCAGCAACCGGAGGAGACGGCTGCAGTTCAGCGCCTCAATGGCGAGACCTGAC	60
DB	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGATTC	120
DB	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnLysProPhe	62
QY	121	GAATGGCGCGCTCGCCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGCCT	180
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	82
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTG	240
DB	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
QY	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAGGTCTCGATCTCAG	300
DB	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	119
QY	301	TCCCAAGACACCAAGACCTCTCCAGGAGAGACCAAGCCACAGCAACAGATAGT	360
DB	120	SerSerArgPro-----GlnAspArg	126
QY	361	CACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTTCAGTTCACCGGTGTGCT	420
DB	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
QY	421	TTCTGGCTTACACGACACGACACTGATGTTGTGCTGTTTCTTCTTACTGACACCAAC	480
DB	147	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	166
QY	481	AACAACGACACGAGTTGATGATTCCTCCCGAGAGATTCAAATTTGGCTGGGAACAGGAG	540
DB	167	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
QY	541	CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAGACGACGAGAACCTTACCATAT	600
DB	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----	199

QY	601	AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660
DB	200	-----Gly	200
QY	661	CAGCACAGCGCAGAGAAACGAGCAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	220
QY	721	AGCGGCTTACCGCGGAGTTCTCGAACCAAGCCTTCAGGTTTCAGCAGACAGACAGATAGT	780
DB	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
QY	781	CAAAACCTTAAGAGCGCAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840
DB	240	ArgLysLeuGlnGlyGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu	259
QY	841	GGCTCAGAACTTTGAGCCCA-----GATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG	891
DB	260	GlyLeuSerValIleSerProThrGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	279
QY	892	GAATACGATCAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	951
DB	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294
QY	952	AGAGCGAGCGGGAATCGTATTGAAGAGAGAGATCTGCACCGCAAGTGTCTAAAGAGACATT	1011
DB	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311
QY	1012	GGTAGAAAACAGATCCCTGACATCTACAACTTCAAGCTGCTTCACTCAAACTGCCAAC	1071
DB	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
QY	1072	GATCTCAACCTTCAATACCTTAGGTGGCTTGGACCTAGTGTGATGATGATGATGATGATGAT	1131
DB	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351
QY	1132	AGGAATGCAATTTGTTCTGCTCCTACTACACCAACGACACACAGCATCATATATCATGTTG	1191
DB	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371
QY	1192	AGGGGACGGCTCAGCTGCAAGTCGTGGACAGCAAGCGCAACAGAGGTGTACGACGAGGAG	1251
DB	372	AsnGlyArgAlaLeuValGlnValValAsnGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	391
QY	1252	CTTCAAGAGGGTCACCGTCTGTGTGGTCCACAGAACTTCGCGCTCGCTGGAAGTCCAG	1311
DB	392	LeuGlnGluGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411
QY	1312	AGCGAACTTCGAATACGTGGCATTCAAGACAGACACTCAAGGCCAGCATAGCCAACTTC	1371
DB	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431
QY	1372	GGCGGTGAAACTCCGCTCATAGTAACCTGCGGAGGAGGTGTTGCCAAATTCATATGGC	1431
DB	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn	451
QY	1432	CTCCAAGGGAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTTCCTGTTTCCA	1491
DB	452	LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro	471
QY	1492	CGGTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
DB	472	ProLysGluSerGlnArgArgValValAla	481

RESULT 10
 US-10-425-114-43865
 ; Sequence 43865, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
NUMBER OF SEQ ID NOS: 73128
SEQUENCE OF SEQ ID NOS: 73128
LENGTH: 488
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700645818_FLI pep
US-10-425-114-43865

Alignment Scores:
Pred. No.: 2e-120 Length: 488
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 15 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-425-114-43865 (1-488)

QY 1 CGGCAGCAACCGGAGGAGACGGTCCAGTTCAGCGCTCAATCGCGAGACCTGAC 60
DB 30 ArgGluGlnProGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 49
QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
DB 50 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 69
QY 121 GAATGGCGCGGCTCGCCCTCTCTCGTGTAGTCTCGCGGCAACGCCCTTGTAGGCT 180
DB 70 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 89
QY 181 TTCTACTCCAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTTG 240
DB 90 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 109
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAGAGCCTCACACAAAGTGTGTGATCTCAG 300
DB 110 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 126
QY 301 TCCCAAGACCAACAGACGCTCTCCAGGAGGAGAACCAAGCCACAGCAACGAGATAGT 360
DB 127 SerSerArgPro-----GlnAspArg 133
QY 361 CACGAGAGTGCACCGTTTCGATAGGGTGATCTCATTCAGTCCACCGGTGTGCT 420
DB 134 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 153
QY 421 TTCTGGCTCTACACGACCAACGACACTGATGTGTGTGCTTTCTTCTTACTGACACCAAC 480
DB 154 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 173
QY 481 AACACGACCAACGAGTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACAGGAG 540
DB 174 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 193
QY 541 CAAGAGTCTTAAGGTACCAACCAACAAAGCAGACGAGCAGACGAGAGAGCTTACCATAT 600
DB 194 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 206
QY 601 AGCCCATACAGCCGCAAAAGTCAGCTTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 207 -----Gly 207
QY 661 CAGCAGCGCCGACAGAACGAGGAGCAAGAGAGAAACGAAGGTGGAACATCTTC 720

DB 208 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 227
QY 721 AGCGCTTTCACCGCGAGTCTCTGGAACAAGCCTTCAGGTTGACGACGACAGATAGTG 780
DB 228 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 246
QY 781 CAAAACCTTAAGAGGCGAGACGAGAGTGAAGAGAGGAGGACCATTTGTACAGTGAAGGA 840
DB 247 ArgLysLeuGlnGlyGluAsnGluGluGluGlyLysGlyAlaIleValThrValLysGly 266
QY 841 GGCCTCAGAATCTTCAGCCCA-----GATAAAGAGACACGCGCCGCAAGAGAG 891
DB 267 GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu 286
QY 892 GAATACGATGAAGATGAATATGAAATACGATGAAGAGGATGAAGAGCGTGGCAGGGAAGC 951
DB 287 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 301
QY 952 AGAGCAGGGGAATGGTATTGAAGAGACGATCTCCACCGCAAGTGCTAAAAGACATT 1011
DB 302 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 318
QY 1012 GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAACTGCCAAC 1071
DB 319 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 338
QY 1072 GATCTCAACCTTCTAATATTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTAC 1131
DB 339 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 358
QY 1132 AGGAATGATTTTGTGCTCCTACCAACACGACGACGATCATATATCGATTG 1191
DB 359 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 378
QY 1192 AGGGACCGGCTCAGTGCAGCTGTGGACACCAACGCAAGAGTGTACACGAGGAG 1251
DB 379 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 398
QY 1252 CTTCAAGAGGGTCAGCTGTGTCGACAGAACTTCGCGCTCGTGGAAAGTCCAG 1311
DB 399 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 418
QY 1312 ACAGAACTTCGAATAGTGCATTCAGACAGACTCAAGCCCGCAGCATACCCACTC 1371
DB 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
QY 1372 GCGGTGAAACTCCGTCATAGATACTCGCGGAGGAGTGGTGCATAATCATATGCG 1431
DB 439 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
QY 1432 CTTCAAAGGAGCAGCAGGAGGAGTGAAGAAACAACCCCTTCAAGTTCTCGTTCCA 1491
DB 459 LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 478
QY 1492 CCGTCTCAGCACTCCCGAGGCGTGTGGCT 1521
DB 479 ProLysGluSerGlnArgArgValValAla 488

RESULT 11
US-10-425-114-44047
Sequence 44047, Application US/10425114
Publication NO. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114

Alignment Scores:	
Pred. No.:	2.01e-120
Score:	1502.00
Percent Similarity:	71.57%
Length:	492
Matches:	297
Conservative:	68

Best Local Similarity:	58.24%	Mismatches:	91
Query Match:	55.82%	Indels:	54
DB:	15	Gaps:	7
US-10-728-323-3 (1-1524) x US-10-425-114-46294 (1-492)			
QY	1	CGGCAGCAACCGGAGGAGACGGCTGCAGCTTCACGCGCTCAATGCGCAGACCTGAC	60
DB	34	ArgGluGlnProGlnGlnAnGluCysGlnIleGlnArgLeuAenAlaLeuLysProAsp	53
QY	61	AATGCGATTGAATCAGAGCGGTTACATTGAGACTTTGGAACCCCAACACCGAGGTTTC	120
DB	54	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAenAsnLysProPhe	73
QY	121	GAATGCGCGGCTCGCCCTCTCTCGCTTAGCTCTCGCGCGCAACGCGCTTCGTAGCCT	180
DB	74	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	93
QY	181	TTCTACTCCAAATGCTCCCGCAGGAGATCTTCATCCAGCAAGGAAGGGATATCTTGGGTTG	240
DB	94	SerTyrThrAenAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet	113
QY	241	ATATTCCCTGGTTGTCTCAGACACTATGAAGAGCTCACACAAAGGTCTCGATCTCAG	300
DB	114	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	130
QY	301	TCCCAAGACACCAAGACGCTCTCCAAGGAGAGACCAAGGCCAACAGCAACGAGATAGT	360
DB	131	SerSerArgPro-----GlnAspArg	137
QY	361	CACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTCCCAACCGGTGTGCT	420
DB	138	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	157
QY	421	TTCTGGCTCTACACGACACGACACTGATGTGTGTCTCTTCTTACTGACACCAAC	480
DB	158	TyrTrpMetTyrAenAenGluAepThrProValValAlaValSerLeuIleAepThrAen	177
QY	481	AACAACGACACCAAGCTTGATAGTTCCTCCCGAGAGATTCAATTGGCTGGGAACACGGAG	540
DB	178	SerPheGlnAenGlnLeuAepGlnMetProArgArgPheTyrLeuAlaGlyAenGlnGlu	197
QY	541	CAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAGCAAGCATAT	600
DB	198	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly-----Gly	210
QY	601	AGCCCATACAGCCCGCAAGTGCAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	211	-----Gly	211
QY	661	CAGCAGCGCGCAGAGAACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB	212	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAenGluGlyGlySerIleLeu	231
QY	721	AGCGGTTTCAGCGCGGAGTCTCTGACACAGCTTCAGGTTTCAGCAGACAGATAGTC	780
DB	232	SerGlyPheAlaProGluPheLeuGluHisAlaPheVal---AspArgGlnIleVal	250
QY	781	CAAAACCTTAAGCGCAGACCGCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840
DB	251	ArgLysLeuGlnGlyGluAenGluGluGluGlyGlyAlaIleValThrValLysGly	270
QY	841	GGCTCAGATCTTGACCCA-----GATAGAAAGAGAGCTGCCGACGAAGAGAG	891
DB	271	GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu	290
QY	892	GATACGATCAAGATCAATATGATACGATGAGAGATAGAGAGCGTGGCAGGGAAGC	951
DB	291	LysProAep-----CysAepGluLysAspLysHisCysGlnSerGlnSer	305
QY	952	AGAGCGAGGGGGAATGGTATTGAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATT	1011
DB	306	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	322
QY	1012	GGTAGAAACAGATCCCTGCATCTACAAACCTCAAGCTGGTTCACCTCAAAATGCCAAC	1071
DB	323	GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrThrAlaThr	342
QY	1072	GATCTCAACCTCTTAATACTTAGTGGCTTAGTGGACCTAGTGGTGAATATGGAATCTCTAC	1131
DB	343	SerLeuAaspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	362
QY	1132	AGGAATGCAATGTTTCTCGCTCACTACAACCAACGACACAGCATCATATCGATTG	1191
DB	363	LysAsnAlaMetPheValProHisTyrAenLeuAenAlaAenSerIleIleTyrAlaLeu	382
QY	1192	AGGGAGCGGCTCACGTGCAAGTCGAGCAGCAACGCGCAACAGAGTGTACGACGAGGAG	1251
DB	383	AsnGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGlyGlu	402
QY	1252	CTTCAAGAGGGTCACGTGCTTGGTGGCCACAGAACTTCGCGCTCGCTGGAAAGTCCCG	1311
DB	403	LeuGlnGluGlyGlnValLeuIleValProGlnAenPheAlaValAlaIleArgSerGln	422
QY	1312	AGCGAAGACTTCGAATACGTGGCATTCAGACAGACTCAAGCCCGCAGCATAGCAACCTC	1371
DB	423	SerAspAenPheGluTyrValSerPheLysThrAenAspArgProSerIleGlyAenLeu	442
QY	1372	GCGGTGAAACTCCGTCTAGATAACCTGCGGAGGAGGTGTTGCCAAATTCATATGGC	1431
DB	443	AlaGlyAlaAenSerLeuAenAlaLeuProGluGluValIleGlnGlnThrPheAen	462
QY	1432	CTCCAAGGGAGCAGGCAAGGAGCTTAAAGAACCAACCCCTTCAAGTTCCTTCGTTC	1491
DB	463	LeuArgGlnGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro	482
QY	1492	CCGTCTCAGCAGCTCTCCGAGGGCTGTGGCT	1521
DB	483	ProLysGluSerGlnArgValValAla	492
RESULT 14			
US-10-425-114-71959			
; Sequence 71959, Application US/10425114			
; Publication No. US2004003488A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(5313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 71959			
; LENGTH: 466			
; TYPE: PRT			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: 700676992_FLI.pep			
US-10-425-114-71959			
Alignment Scores:			
Pred. No.:	3,23e-120	Length:	466
Score:	1499.50	Matches:	297
Percent Similarity:	72.02%	Conservative:	71
Best Local Similarity:	58.12%	Mismatches:	94
Query Match:	55.72%	Indels:	49
DB:	15	Gaps:	6
US-10-728-323-3 (1-1524) x US-10-425-114-71959 (1-466)			
QY	1	CGGCAGCAACCGGAGGAGAGAACGGTTCAGCGCTCAATGCGCAGACCTGAC	60

```
Db 1 ArgGluGlnAlaGlnGlnAenGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
  |||::||| :::::|||| ||||| |||::||| ||||| ::|
Qy 61 AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 121 GAATGGCGCGGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGGCCT 180
  ::|::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 41 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 60
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTG 240
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 61 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 80
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTCTGATCTCAG 300
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 81 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 100
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGAGACCAAGCCACAGCAACGAGATAGT 360
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 101 SerGlnArgPro-----GlnAspArg 107
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 361 CACAGAAGTGTACCGTTTCGATGAGGGTGTATCTCATTCAGTTCCTCCACCGGTGTGCT 420
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 108 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 421 TTTCTGGCTTACACGACCAACGACACTGATGTGTGCTGCTTCTTCTACTGACACCAAC 480
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn 147
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 481 AACAAACACACACAGCTTGATAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGAG 540
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 148 SerLeuGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 167
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 541 CAAGAGTCTTAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATAT 600
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 168 GlnGluPheLeuLysTyrGlnGlnGln----- 176
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 601 AGCCCATACCGCGCAAGCTAGCTAGACAGAGAGCGGTGAATTTAGCCCTCGAGGA 660
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 177 -----GlnGlnGlyGlySerGlnSerGlnLysGly 186
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 661 CAGCACAGCCGAGAGAAAGCAGCAGCAGCAAGAAAGAAAGAAAGCTGGAACATCTTC 720
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 187 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 199
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 721 AGCGGTTCACGCGGAGTTCCTGGAAACAGCCCTTCAGGTTCAACACAGACAGATAGTG 780
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal--AsnMetGlnIleVal 218
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 781 CAAACCTAAGAGCGCAGCGAGAGTGAAGAGAGGAGCCATTCGACAGTGAAGGGA 840
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 219 ArgAsnLeuGlnGlyGluAsnGlnGluGluAspSerGlyAlaIleValThrValLysGly 238
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 841 GGCTCTCAGAACTTGAGCCAGATAGAAGAGACGTCGCCAGCAAGAGAGAAATACGAT 900
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 901 GAAGATGAA-----TATGAATACGATGAAGAGATAGAAGCGGTGGCGGGA 948
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 259 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 949 AGCAGAGGCGGGGATGTATTGAAGAGACCATCTGCACCCCAAGTGTCTAAAGAAC 1008
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1009 ATTTGGTAGAAACAGATCCCCCTGACATCTACAACCCCTCAAGCTGGTTCATCTCAAAACTGCC 1068
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 296 IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrThrAla 315
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1069 AACGATCTCAACCTTCTAATACTAGTGGCTGGACCTAGTCTGGAATATGAATCTC 1128
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
```

```
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1129 TACAGAAATGCAATGTTTGTCTCACTACAAACCAACGACGACATCATATATCGA 1188
  ::|::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 355
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1189 TTGAGGGGACGGGCTCAGCTGCAAGTCGTGGACCAACGCGCAACAGAGTGTACGACGAG 1248
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1249 GAGCTTCAAGAGGGTCAGCTGCTTGTGTGTCACAGAACTTCGCGTGTGCGAAGTCC 1308
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 376 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 395
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1309 CAGACGGAACCTCGAATACGTGCACTTCAAGACAGACTCAAGCCACGACATAGCCAAC 1368
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 396 GlnSerAspAsnPheIleTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1369 CTCGCGGTGAAAACCTCCGTCATAGATAACCTGCCGAGGAGGTGGTTCGCAAAATTCATAT 1428
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 416 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 435
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1429 GGCCTCCAAAGGAGCAGGACGAGGAGCTTAAGAACAAACCCCTTCAAGTCTTCGTT 1488
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 455
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 1489 CCACCGTCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 456 ProProGlnGluSerGlnArgArgAlaValAla 466
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
```

RESULT 15

```
US-10-425-114-71973
; Sequence 71973, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71973
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700677092_FLI.pep
US-10-425-114-71973
```

Alignment Scores:

Pred. No.:	3-23e-120	Length:	466
Score:	1499.50	Matches:	297
Percent Similarity:	72.02%	Conservative:	71
Best Local Similarity:	58.12%	Mismatches:	94
Query Match:	55.72%	Indels:	49
DB:	15	Gaps:	6

US-10-728-323-3 (1-1524) x US-10-425-114-71973 (1-466)

```
Qy 1 CGGCGAACCAGGAGAGAACCGCTGCCAGTTCACGCGCTCAATCCGCGAGACCTGAC 60
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 1 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Qy 61 AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
  |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
```


This Page Blank (uspto)